Analyse des corrélations spatiales des quasars et implications cosmologiques avec le multi-spectrographe SDSS-IV eBOSS

Pauline Zarrouk

NNT : 2018SACLS297

Thèse de doctorat de l'Université Paris-Saclay préparée à l'Université Paris-Sud

CEA/DRF/Irfu/DPhP
École doctorale n°576
Particules, Hadrons, Énergie, Noyau, Instrumentation, Imagerie, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat: Astroparticules et Cosmologie
par

Pauline Zarrouk

Clustering Analysis in Configuration Space and Cosmological Implications of the SDSS-IV eBOSS Quasar Sample

Thèse présentée et soutenue à l'amphithéâtre Bloch, IPhT, Gif-sur-Yvette le 1er octobre 2018.
Composition du jury :

Mr Shaun Cole	Professeur Durham University, ICC	Rapporteur
Mr Francis Bernardeau	Directeur de recherche Sorbonne Université, IAP	Rapporteur
Mme Françoise Combes	Professeur Observatoire de Paris, LERMA	Présidente du jury
Mr Jean-Paul Kneib	Professeur EPFL, LASTRO	Examinateur
Mr Filippo Vernizzi	Chargé de recherche Université Paris-Saclay, CEA IPhT	Examinateur
Mr Martin Kilbinger	Directeur de recherche Université Paris Saclay, CEA Irfu DAp	Examinateur
Mr Etienne Burtin	Directeur de recherche Université Paris-Saclay, CEA Irfu DPhP	Directeur de thèse

Préambule

This PhD manuscript is written as completion of my work at the Department of Particle Physics, Institut de Recherche sur les lois Fondamentales de l'Univers (IRFU) at CEA-Saclay from 2015 to 2018. These three years as a PhD student were more than just preparing an additional and last diploma. It was above all an amazing human adventure where I feel very lucky to meet lots of wonderful people who enriched me.

How to read this thesis

I decided to write my manuscript in English as research nowadays is made in English and I wanted to give the possibility for potential non-French PhD students and post-doctoral researchers to enjoy and use my work. In particular, I wanted my first chapter to provide a comprehensive and deep introduction to modern cosmology both presenting the key theoretical developments and the status of observations with the hope that it will be useful for master students and future PhD students.

However, I would also like my relatives, my friends and any curious person to be able to understand the stakes of my research subject at least. Therefore, this preamble first proposes a short abstract in French that I tried to make more accessible to non-scientific people. To all people who are curious about the universe in which we live, I wish you to enjoy the reading!

Résumé vulgarisé de mon travail de thèse

A tous les curieux de l'univers dans lequel nous vivons,

Notre univers a une histoire.
"D'où venons-nous et comment en sommes-nous venus à exister ?" Telles sont les questions posées par l'astrophysicien Hubert Reeves dans *Chronique des atomes et des galaxies*, le premier livre scientifique sur l'univers que j'ai lu alors que j'étais en Première S au lycée. J'ai découvert que la cosmologie pouvait apporter des réponses scientifiques à ces questions.

Notre univers est immense.
"Terre, planète bleue, où des astronomes exaltés capturent la lumière des étoiles aux confins de l'espace." *Hubert Reeves*
Pour illustrer les dimensions vertigineuses de notre univers, imaginons que le système solaire ait la taille d'un grain de sable. Alors une galaxie aurait la taille du rayon de la Terre et la distance moyenne entre deux galaxies correspondrait à la distance Terre - Lune !

La cosmologie a pour but de retracer l'histoire de l'univers en s'appuyant à la fois sur des modèles mathématiques et sur des observations. Dès les années 1920, les observations du mouvement des galaxies, ces regroupements d'étoiles attirées sous l'effet de leur masse, ont montré que plus les galaxies sont lointaines, plus elles s'éloignent vite les unes des autres : l'univers est en expansion. A la fin des années 90, d'autres observations ont révélé que les distances entre les galaxies augmentent plus rapidement lors des derniers 6 milliards d'années de l'univers, lorsque les chercheurs estiment que notre univers est âgé de 13,8 milliards d'années. Cette découverte capitale de l'accélération de l'expansion de l'univers est l'une des plus grandes énigmes de la cosmologie. Afin d'expliquer les observations, le modèle actuel suppose qu'il existe une composante exotique, appelée "énergie noire", qui domine aujourd'hui le contenu énergétique de l'univers. A ce jour, aucune théorie ne permet d'expliquer de manière satisfaisante l'origine de cette accélération.

Dans le cadre de mon doctorat, je fais partie de l'un des plus ambitieux programmes d'observation du ciel porté par la collaboration Sloan Digital Sky Survey (SDSS) qui utilise un télescope aux états-Unis en opération depuis les années 2000 et qui a déjà mesuré la position dans le ciel de plus d'un million de galaxies. En comptant le nombre de fois où deux galaxies sont séparées par une distance donnée, on reconstruit la distribution des corrélations spatiales des structures cosmiques de l'univers. La mesure de cette distribution au moyen de grands relevés comme le SDSS a ouvert une nouvelle piste pour explorer la nature de l'énergie noire grâce à l'exploitation d'une nouvelle sonde : les oscillations acoustiques de baryons (BAO). Il s'agit d'ondes de pression qui se sont propagées pendant les 380 000 premières années de l'univers et qui présentent une signature caractéristique dans la carte des structures de l'univers : deux galaxies sont préférentiellement séparées de 500 millions d'année-lumière. Cette distance caractéristique (ou échelle BAO) est utilisée comme un étalon de distance pour mesurer l'évolution des distances dans l'univers, autrement dit son taux d'expansion. Pour étendre les mesures faites jusqu'à présent grâce aux galaxies et être capable de sonder l'univers tel qu'il était il y a plus de 6 milliards d'années, il faut pouvoir observer des astres très brillants et très lointains. Les quasars sont justement les candidats idéaux pour établir la carte la plus aboutie des structures de l'univers. Il s'agit de trous noirs super-massifs situés au centre de certaines galaxies qui émettent une très forte lumière provenant du disque d'accrétion de matière situé autour du trou noir central. Ces phares cosmiques sont précisément les objets que j'étudie dans ma thèse et que le programme SDSS-IV eBOSS a observé pendant deux ans pour collecter la position de presque 150 000 d'entre eux. Mon sujet de thèse porte sur l'étude de la distribution des corrélations spatiales des quasars de eBOSS pour contraindre l'énergie noire à une époque de l'univers pratiquement inexplorée à ce jour. Il s'agit d'une analyse clé de ma collaboration dans laquelle j'ai joué un rôle de premier plan. Une première partie de mon travail de recherche a constitué à mesurer l'échelle BAO à partir de l'échantillon de quasars observés par eBOSS et a donné lieu à une publication dans la revue scientifique d'astronomie Monthly Notices of the Royal Astronomical Society (MNRAS). « Il y a le mètre pour des petits échelles de distances, le kilomètre ou le mile pour les distances entre les villes, et nous avons l'échelle BAO pour des distances entre galaxies et quasars en cosmologie. » ai-je expliqué dans le communiqué de presse de ma collaboration au printemps 2017 (repris pour le CEA-Irfu en français ici) où, bien que la collaboration compte presque 150 personnes et que le milieu soit très compétitif, je suis la seule doctorante à avoir été citée en reconnaissance pour ma contribution significative dans cette analyse.

Plusieurs pistes sont envisagées pour comprendre le phénomène d'accélération de l'ex-

pansion de l'univers. En plus de mesurer précisément l'évolution des distances pour contraindre les propriétés de l'énergie noire, il est possible de tester la validité de la théorie de la gravitation aux échelles des structures cosmiques. L'interaction gravitationnelle qui est responsable de la formation des structures qui s'attirent sous l'effet de leur masse, est décrite par la théorie de la relativité générale d'Einstein qui est testée de manière très précise à l'échelle de notre système solaire. Mais nous supposons que la théorie est valable aux échelles de distances entre les galaxies, autrement dit à des distances beaucoup beaucoup plus grandes. Pour reprendre l'analogie entre la taille de notre système solaire et le grain de sable, nous savons actuellement confirmer la validité de la théorie d'Einstein pour des distances qui correspondent au grain de sable et nous supposons qu'elle est valable pour des distances qui correspondent à la distance entre la Terre et la Lune. Il est donc possible qu'à ces échelles cosmologiques, des modifications soient à apporter à la théorie de la relativité générale. Ainsi, mon travail de recherche comprend un deuxième volet plus novateur qui consiste à mesurer la quantité de galaxies ou quasars qui se forment à une période donnée de l'univers. On appelle ce paramètre cosmologique le taux de croissance des structures cosmiques et sa valeur est prédite par le modèle actuel de la cosmologie. Nous avons mené une analyse rigoureuse des effets pouvant altérer la mesure qui a permis de déterminer la croissance des structures pour la première fois à partir de cet échantillon de quasars. Nous avons ainsi confirmé la validité du modèle actuel de la cosmologie basé sur la relativité générale pour une époque de l'univers quasiment inexplorée à ce jour. Cette étude a fait l'objet d'un article 1er auteur dans la revue MNRAS et j'ai accompagné cette publication d'un communiqué du CEA-Irfu en février.

Par ailleurs, aucune expertise dans la mesure du taux de croissance des structures n'existait dans mon groupe de recherche au CEA-Saclay, et était peu développée en France. Mon travail de recherche est d'autant plus précurseur qu'il ouvrira la voie pour les futurs grands relevés, comme le prometteur programme DESI qui s'inscrit dans la continuité du relevé SDSS ou encore le satellite européen Euclid. Le défi majeur de ces programmes dédiés à l'énergie noire sera de cartographier les structures de l'univers pour mesurer les paramètres cosmologiques avec un gain en précision d'un ordre de grandeur par rapport aux contraintes actuelles. De plus, les mesures de croissance des structures à grand redshift, comme la première mesure que j'ai effectuée avec l'échantillon quasars de eBOSS vont apporter des nouvelles contraintes sur les modèles de gravité alternative, des contraintes qui sont complémentaires de celles qui sont en train d'être apportées par les ondes gravitationnelles et qui s'annoncent aussi très prometteuses.

Remerciements, Acknowledgments

J'aimerais d'abord exprimer toute ma gratitude envers mon directeur de thèse, Etienne Burtin. Je te remercie sincèrement de m'avoir accordé ta confiance lorsque je suis venue pour mon stage de master 2 et un sujet de thèse. J'ai tout de suit été séduite par ton enthousiasme et tes efforts de pédagogie, en plus du coup de cœur évident pour le sujet et le groupe. Je te remercie d'avoir toujours été très disponible et accessible, de m'avoir considérée dès mon stage comme une collègue plus qu'une étudiante et de m'avoir toujours encouragée dans toutes mes initiatives, même en dehors du strict cadre de la thèse pour ma participation à de nombreux événements. Faire de la recherche c'est aller au-delà de ce que l'on sait, et toi tu m'as appris à chercher, à faire parler une figure jusqu'à ce qu'elle nous révèle la prochaine étape, à rendre accessible l'intuition scientifique. J'ai beaucoup d'admiration pour le chercheur que tu es, j'en ai tout autant pour l'être humain. J'espère avoir encore de nombreuses opportunités de travailler avec toi, mais aussi de discuter entre "hommes révoltés" autour d'une IPA...

Je remercie chaleureusement Anne Isabelle Etienvre, anciennement chef du SPP lorsque j'ai commencé ma thèse, maintenant chef de l'Irfu, et qui dès mon stage nous a soutenu, Etienne et moi, dans notre quête d'un financement. Je remercie également le labex d'excellence P2IO (Physique des 2 Infinis et des Origines) de l'Université Paris-Saclay qui m'a accordée une bourse P2IO puis le prix de thèse P2IO 2018.

Je remercie le groupe de cosmologie du DPhP pour son accueil chaleureux, j'y ai découvert une famille professionnelle aux multiples personnalités et qui constitue l'une des richesses du groupe que j'ai le plus appréciée. Je remercie Christophe Y. pour son franc-parler et son leadership, Nathalie pour sa capacité à mêler excellence scientifique et vulgarisatrice passionnée, Vanina pour nos discussions toujours enrichissantes sur les modèles de gravité modifiée et sur la société, Christophe M. pour avoir partagé sa passion de l'instrumentation et m'avoir permis de participer à une manipulation en salle blanche pour DESI, Jean-Baptiste pour nos échanges sur *Planck* et pour sa motivation à organiser ou participer à des événements de toute sorte, Jean-Marc, Jim et Eric pour les nombreuses discussions également passionnantes. Je remercie également Roy pour ses anecdotes sur une époque de sa vie que j'aurais aimé connaitre et qui fait partie de ces personnes qui m'ôte la peur de vieillir. Je remercie également les anciens doctorants, Pierre qui m'a parlé pour la première fois du groupe cosmo alors que j'effectuais un stage de L3 en physique des particules au SPP, Julien mon irremplaçable collègue de bureau, Hélion, Clément mon binôme de promo mais aussi les doctorants actuels, Charles-Antoine, Arnaud, Solène ma co-bureau pendant ma dernière année, Thomas, Emmanuel, et Richard, le nouveau doctorant d'Etienne. Je lègue mon canapé aux occupants du bureau 36A tant que ce bureau demeure le bureau d'un étudiant d'Etienne et que le groupe continue d'en profiter. Je vous remercie tous pour votre bonne ambiance, je compte sur la Junior Cosmo Team pour perpétuer certaines traditions comme les goûters et sur les permanents pour continuer de participer et d'encourager ces moments en dehors du strict cadre du travail qui, pour

moi, ont été essentiels à l'entretien de mon enthousiasme et de ma motivation. Je remercie également tous les autres collègues que j'ai pu rencontrés et avec qui j'ai eu la chance de discuter, voir de collaborer. Un merci tout particulier à Laurent Chevalier mon parrain au DPhP, Filippo Vernizzi de l'IPhT, Martin Kilbinger du DAp, et à toutes les personnes qui oeuvrent au rapprochement des équipes de recherche dans la même thématique au sein du CEA. J'espère que le groupe de cosmologie du DPhP va continuer de grandir et de s'imposer comme un groupe incontournable dans l'analyse et l'exploitation des données spectroscopiques, complémentaires des données photométriques davantage utilisées au DAp. Je suis très fière de savoir que l'Irfu est engagé dans les deux prochains programmes d'observation, DESI au sol et Euclid dans l'espace, j'espère que les cosmologistes du DPhP, les astrophysiciens du DAp et les théoriciens de l'IPhT continueront de s'enrichir mutuellement de leurs différentes expertises.

Je tiens également à remercier Gautier Hamel de Monchenault et Georges Vasseur, respectivement chef et chef-adjoint du DPhP pour l'intérêt qu'ils ont porté à mes travaux de recherche et leur soutien dans ma participation aux réunions de collaboration, conférences et ateliers de travail internationaux auxquels je souhaitais assister ou pour lesquels j'étais invitée. De manière générale, je remercie la direction de l'Irfu qui m'a apportée un fort soutien et une nette reconnaissance de mes travaux en me permettant de partir deux mois au Centre d'astrophysique de Harvard pour travailler avec Daniel Eisenstein, expert de renommée internationale. Je remercie également Martine Oger, secrétaire exemplaire du DPhP sans qui les démarches administratives dans un laboratoire de recherche, qui plus est au CEA, auraient pu être beaucoup plus cauchemardesques.

I am also very grateful to my collaboration SDSS-IV eBOSS and in particular to my galaxy quasar clustering working group. I won't never forget the first time I presented myself and my research at the eBOSS collaboration meeting at Lausanne in February 2016. It marked the beginning of intense and constructive exchanges within my working group where I felt encouraged and supported all along my thesis work. In particular, I acknowledge fruitful discussions with Will Percival, Jean-Paul Kneib, Ashley Ross, Kyle Dawson, Héctor Gil-Marin, GongBo-Zhao, Anand Raichoor, Johan Comparat, Rita Tojeiro, Julian Bautista. A special warm thank you to Daniel Eisenstein who welcomed me during 2 months in his institute at Harvard while I was starting my 3rd and last year. It was a quite difficult period of my thesis and I thank you very much for guiding me in my research career and giving me the opportunity to work with you and benefit from your deep experience. You are all people I am glad and proud to collaborate with and I hope it will continue. I am also very grateful to all the organizers of the different conferences and workshop I attended. In particular I acknowledge the excellent program of the workshop Consistency of cosmological datasets at KICC where I am very grateful for being invited to give a talk on behalf of my working group. I also thank the organizers of the Rencontres de Moriond for the quality of the program and the unique spirit associated to this week in the Alps that I hope I will be able to feel again. I am also very grateful to the organizers of the French Euclid summer school that I attended in July 2017. It was really nice to meet other researchers who work in the same field and enrich my existing but mosly international network with potential colleagues in France. A special thank you to Sandrine Codis and Francis Bernardeau from IAP, Sylvain de la Torre and Eric Jullo from LAM and Alain Blanchard from IRAP. You are definitely people I will be glad to collaborate with.

Regarding the examination committee of my PhD defense, I am very proud I was able to gather as many experts in my jury. I wanted to have an expert of each sub-topic I enjoyed during my PhD work. I thank warmly my two reviewers, Shaun Cole and Francis

Bernardeau for their attentive reading and fruitful feedbacks, I look forward to working with you at Durham, Shaun and I have no doubt I will learn a lot in your group. Francis, your rewiew on perturbation theory is still my bedside reading, I am glad we met at the Euclid summer school as it gave me the opportunity to start a project with you for Euclid and I hope it will continue. I thank also the other members of my committee : Françoise Combes whose lectures on quasars at Collège de France that I attended during my first year of PhD were very interesting, Jean-Paul Kneib who is the Principal Investigator of our eBOSS collaboration and who I was glad to met during the first collaboration meeting I attended, Filippo Vernizzi for fruitful discussion on effective field theory and modified gravity models, Martin Kilbinger whose lectures on weak lensing at the Euclid summer school were very instructive, and of course Etienne my PhD advisor.

Au cours de ces 3 ans, j'ai eu et je me suis donnée l'opportunité de mener d'autres activités en dehors de ma thèse. Je remercie Yan Picard avec qui j'ai effectué une mission d'enseignement à l'Université Paris-Sud et qui m'a proposé un service exceptionnel, avec notamment la possibilité d'assurer en toute autonomie une partie des cours de l'option physique en PCSO. Je remercie également l'équipe Astronomie du Palais de la découverte, en particulier Sébastien Fontaine son responsable, et Marc Goutaudier pour leurs régulières invitations à donner des conférences publiques, ainsi que toutes les personnes qui m'ont invitée pour des manifestations grand public, que ce soit dans le cadre de la Fête de la Science, de Pint of Science, du festival Pariscience ... Je tiens tout particulièrement à remercier Bruno Monflier, président de l'association Ferme des Etoiles et du festival d'astronomie de Fleurance, ainsi que Nathalie de mon groupe et l'ensemble du comité scientifique du festival pour m'avoir donné l'opportunité d'être marraine et conférencière de l'édition 2017 du festival aux côtés d'Hubert Reeves. Cet été 2017, un premier rêve s'est concrétisé et je veux continuer de me nourrir de toutes ces rencontres chaleureuses avec le grand public. Je remercie également l'association Les amis de la Cité de l'Espace avec son président Michel Bouffard pour m'avoir décernée le Prix des Jeunes 2017. Je suis également porteuse d'un projet de jeu de plateau sur les exoplanètes qui a été sélectionné par l'appel à projet Coup de pouce de la Diagonale Paris-Saclay que je remercie beaucoup pour leur soutien qui nous permet de distribuer gratuitement des exemplaires dans les établissements scolaires, lieux de diffusion scientifique et laboratoires de recherche. Le jeu a été crée par Alexandre Bordas, Dimitri Chuard et moi-même dans le cadre d'un enseignement universitaire de vulgarisation de la physique proposée par Julien Bobroff et Frédéric Bouquet que je remercie vivement. Le projet bénéficie également du soutien du service communication de l'Irfu, et je tiens à remercier Philippe Briet, Sophie Kerhoas-Cavata et Estelle Lemaitre pour m'avoir fait confiance et accompagné dans cette nouvelle aventure ! Je remercie également Loick Drappier et Caroline Gaudillère de l'entreprise de graphisme Idées fraiches et Héger Benaissa pour leur participation enthousiaste à la réalisation de ce projet.

Je suis également l'une des 30 lauréates de la bourse L'Oréal-Unesco Pour les Femmes et la Science Edition 2018, je remercie l'ensemble des personnes responsables du programme de ces bourses de m'avoir sélectionnée, j'espère prouver qu'une carrière scientifique de haut niveau est compatible avec une vie de femme et augmenter ainsi la portée de mon message grâce au réseau du programme international 'For Women in Science'.

Enfin, je remercie Marie-Hélène Schune pour m'avoir parlé du sujet de thèse proposé par Etienne alors que j'étais en master 2 NPAC. Je remercie également Yves Charon pour avoir proposé une sensibilisation à l'insertion professionnelle des docteurs en master 2, cette réflexion ayant conduit à la réalisation d'un état des lieux sur le dispositif de mission doctorale de conseil en entreprise méconnu en France alors que je suis convaincue qu'il

est un des moyens les plus souple et les plus performant à l'heure actuelle pour mettre les compétences uniques des doctorants formés par la recherche au service des entreprises. Je suis ainsi devenue la présidente de l'association Synapse qui œuvre à faire connaitre et mettre en place les missions doctorales de conseil en entreprise au niveau de l'université Paris-Saclay. Je remercie les membres de cette association pour leur engagement, j'espère que nous aurons l'occasion de continuer les actions de Synapse. J'ai également eu l'opportunité de devenir membre du comité d'organisation des Rencontres Physique-Entreprise-Recherche organisées en mars 2017 par la Société Française de Physique, dont j'ai également rejoint le conseil administratif. Je remercie en particulier Michel Spiro, son ancien président, Jean-Jacques Bénattar, Eric Perez, Mayline Gautié, la Commission Jeunes et l'ensemble des personnes qui ont participé à l'organisation des rencontres et qui font vivre la SFP. J'espère que nous parviendrons à mieux valoriser le doctorat en France et à faire connaitre davantage les docteurs auprès des entreprises et institutions gouvernementales.

Pour finir, j'aimerais remercier spécifiquement des personnes que j'ai eu la chance de rencontrer et qui ont suscité ma passion pour la cosmologie. Je pense bien sûr à Hubert Reeves que j'ai eu l'immense plaisir de côtoyer pendant plusieurs jours au festival d'astronomie de Fleurance, merci pour votre générosité, votre accessibilité et votre humilité. "Il y a quelque chose à faire des quelques décennies que la nature nous accorde, prendre résolument et sans faillir le parti d'embellir la réalité". Je vais continuer de m'efforcer de suivre au mieux votre conseil. J'ai eu la chance de rencontrer d'autres astrophysiciens qui sont sortis de leur laboratoire pour venir à la rencontre du grand public, je remercie en particulier Nathalie Palanque-Delabrouille, Jean-Philippe Uzan, Hervé Dole, Roland Lehoucq et André Brahic. Toutefois mon intérêt pour la cosmologie et ma motivation pour poursuivre une carrière scientifique ne m'auraient pas suffit pour arriver là où je suis maintenant, si 3 personnes, 3 professeurs n'avaient pas été là. Je remercie très vivement Christophe Jorssen, mon professeur de physique-chimie au lycée, passionné d'astrophysique, et dont les conseils m'ont définitivement convaincue que les études scientifiques s'annonçaient passionnantes, et surtout à ma portée. Je remercie également Dany Pellequer et Vincent Sauzéat, mes professeurs d'espagnol et de musique respectivement pour leur soutien indéfectible. Enfin, mes dernières pensées et remerciements vont à ma famille dont les quelques mots qui vont suivre ne pourront suffire à exprimer tout ce que je lui dois. Je sais que mon exigence et mon savoir-vivre sont familiaux, ma créativité maternelle et ma persévérance paternelle. Je sais aussi qu'une carrière dans la recherche fondamentale en cosmologie peut effrayer beaucoup de parents, les miens l'ont été puis j'ai avancé dans mes études, commencé ma thèse et ils ont réalisé que j'étais exactement là où je voulais être. Je remercie donc aussi mes proches pour avoir osé me faire part de leurs inquiétudes, elles n'ont fait que renforcer ma motivation. La famille Zarrouk s'est désormais enrichie d'un nouveau docteur, en cosmologie cette fois ! Et enfin, je remercie mon compagnon, Alexandre Bordas, qui partage ma vie depuis 6 ans maintenant et qui m'aime tout simplement comme j'aime être aimée. Je terminerai avec cette citation d'Alfred de Musset que tu connais bien : "Je ne sais pas où va mon chemin mais je marche mieux quand ma main serre la tienne." C'est peut-être la seule certitude que j'ai, mais elle est indéniable et elle me suffit à avancer à tes côtés.

Paris, 15 novembre 2018.

Table des matières

Chapitre 1

Introduction to modern cosmology

Contents

1.1 The standard cosmological model

Modern cosmology began a century ago with the discovery of the expansion of the universe, both as a theoretical solution of the equations that govern the dynamics of the

universe and as observational evidence to explain that galaxies move away from each other. Nowadays, it is well described within the framework of the standard cosmological model which follows the expansion of the universe over 13.8 billion years starting with a 'hot Big Bang'. The large-scale dynamics of the universe is described by the theory of General Relativity (GR) applied to an homogeneous and isotropic universe whose present day energy content is illustrated in figure 1.1 and is made-up of :

— **Baryonic matter** represents a small fraction of the total matter content. It is also called *ordinary matter* as it corresponds to galaxies, stars, gas, planets, us, ...

— **Cold dark matter** is the dominant matter component and is assumed to be non-relativistic and pressureless. Its microscopic nature is still unknown with masses ranging from 10^{-22}eV to 10^{67}eV. Candidates for dark matter can be studied in particle physics with the Large Hadron Collider (LHC) [1] and with cosmology through its effects on the formation of structures.

— **Radiation** is a relativistic component that includes photons and neutrinos while they were relativistic. Its contribution is negligible today but in the early universe, when temperatures were higher, it was dominant and coupled with matter.

— **Dark energy** dominates the energy content today and is responsible for the late-time acceleration of the expansion of the universe. In the standard model, dark energy is a fluid with constant properties in time and space that can be characterized by a **cosmological constant** which enters the equations of general relativity.

Given the energy content of the universe, this standard model is also called the ΛCDM model. The universe cooled down due to the expansion and the densities of each component evolved with time so that the universe underwent different eras, first dominated by radiation, then matter and eventually by dark energy. The theoretical predictions of the ΛCDM model match many different observations from the photon radiation emitted 380,000 years after the Big Bang to today's large-scale structures of the universe. But the fact that it relies on a *dark sector* (dark matter and dark energy) that represents almost 95% of the current energy content of the universe reminds us that a large part of discoveries about the universe are still to be made, and as a PhD student it is a very exciting time to start learning about modern cosmology.

This section is organized as follows. Section 1.1.1 describes the most important notions and results of GR (section 1.1.1.1) to focus on its application in cosmology (section 1.1.1.3) in order to obtain a system of equations that governs the evolution of the universe and its content. The resolution of these equations in the 1920's led to a time-evolving solution, meaning that the universe is not static, contrary to the general viewpoint at that time. We will see in section 1.1.1.2 that the idea according to which the universe is expanding took about thirty five years to be widely accepted by the scientific community. It led to the current scenario of a *hot Big Bang* model that was originally based on three observations : the expansion of the universe, the light element abundances which are in agreement with the Big Bang nucleosynthesis (BBN), and the black-body radiation left over from the first 380,000 years, the cosmic microwave background (CMB). In the last three sections, we describe the unknowns of this standard model : the origin of primordial fluctuations (which could emerge from quantum fluctuations during the inflation) in section 1.1.3.1, the cold dark matter that plays a crucial role in the formation of structures (section 1.1.4) and the late-time acceleration of the expansion in section 1.1.5 whose mechanism remains

1. https://home.cern/topics/large-hadron-collider

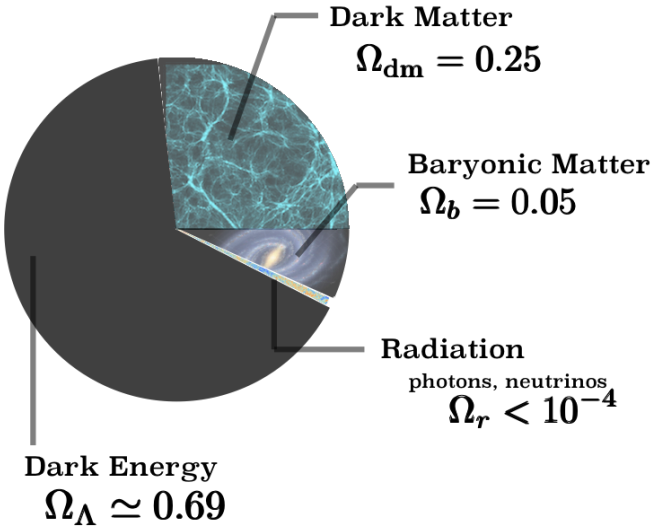


FIGURE 1.1 – Present day energy content of the universe : matter represents ∼30% and is composed of ordinary baryonic matter (∼5%) and a cold dark matter (∼25%) whose nature remains unknown, radiation (photons and neutrinos) is negligible and dark energy, a mysterious fluid which is responsible for the cosmic acceleration, is now dominated the energy content (∼70%).

one of the most important questions of modern cosmology.

This section is heavily influenced by the lectures on cosmology during my master 2 NPAC program [2], by textbooks [1, 2, 3, 4, 5, 6]. The mathematical framework to describe the statistical properties of Gaussian random fields and their application for cosmology in section 1.1.3.2 are also inspired by Benjamin Wandelt's lectures on Cosmostatistics and Large Surveys.

1.1.1 The geometry of the universe

General relativity (GR) provides a geometrical explanation of the gravitational interaction which is no longer considered as an instantaneous force between two massive particles as Newton introduced three centuries before. It is worth noting that most cosmology can be learned without a deep knowledge of the mathematical tools that are associated with the description of a curved space-time. One must be familiar with the concept of a metric, understand geodesics and be able to apply Einstein's equations to the metric that describes an isotropic and homogeneous universe. Appling the formalism of GR to the universe will then relate the parameters of the metric to the densities in the universe.

General relativity relies on the same assumptions that special relativity which has been developed to explain situations where objects can move with a velocity close to the speed of light. If we assume the *universality of the speed of light*, the simple addition of velocities does not work for objects moving at the speed of light in a frame which is itself moving. So in 1905, Einstein proposed the theory of special relativity where he introduced the concept of *space-time invariance* which led to a relativistic description with a modified addition of velocities that works for frames moving at a constant speed in a straight line (such frames

2. https://npac.lal.in2p3.fr/home/

are called inertial). Then, he extended his formalism to accelerated frames by assuming that uniform gravitational fields are equivalent to frames that accelerate uniformly relative to inertial frames. This assumption is called the **Equivalence Principle** and it means that the gravitational mass m_g (the one in the definition of the gravitational interaction) is equal to the inertial mass m_i (the one associated with the Newton's second law, $F = m_i a$ where F is the total force and a the acceleration vector). This principle is heavily tested on Earth and in the Solar System as it is a foundation of general relativity and any deviation in the equivalence $m_i = m_g$ would imply collapse of the theory. Among the latest experimental tests of GR, the recently launched MICROSCOPE satellite confirms the validity of the Equivalence Principle at the 10^{-15} precision level [7].

1.1.1.1 Space and time in General Relativity

Classical vs Relativistic description In classical physics, if one observer measures the separation to be δL, then all observers measure the same δL regardless of how they are moving, and the same consideration is valid for the passage of time : the separation in time between two events is the same for all observers. In special and general relativity (SR and GR), neither spatial separations nor time intervals are invariant but instead there is a combined space-time interval δs which is invariant :

$$(\delta s)^2 = (c\delta t)^2 - (\delta x)^2 - (\delta y)^2 - (\delta z)^2 \tag{1.1}$$

where c is the speed of light in a vacuum. The coefficients in the right-hand side (here, $+1, -1, -1, -1$) represent the *metric coefficients* of the SR metric tensor. The metric tensor and the coordinate system fully describes the geometry of space-time. The latter is now defined by $\mathbf{x} = (x^0, x^1, x^2, x^3)$ where x^0 represents the time coordinate and $x^{1,2,3}$ the spatial coordinates (the Einstein notation for the elements of $\mathbf{x}$ is x^μ). These coordinates define an event in SR and GR. The metric tensor $g_{\mu\nu}(\mathbf{x})$ expresses the distance between two events in space-time and another way of formulating the distance ds between the points $\mathbf{x}$ and $\mathbf{x} + d\mathbf{x}$ is given by :

$$ds^2 = g_{\mu\nu}(\mathbf{x})\, dx^\mu\, dx^\nu \tag{1.2}$$

By construction the metric tensor is symmetric ($g_{\mu\nu} = g_{\nu\mu}$) and special relativity is described by a Minkowski space-time with the metric $g_{\mu\nu} = \eta_{\mu\nu}$, where

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Geodesic equation In Minkovski space (3D space with no curvature), particles travel in straight lines unless they are acted on by a force. In more general space-times, and in particular in a curved space-time, the paths of particles are more complicated and in absence of any forces, they are called *geodesics*. To express this in equations, one need to generalize Newton's law with no forces, i.e. $d^2x^i/dt^2 = 0$, to the expanding universe. This derivation is not the purpose of this section, we will just apply the Equivalence Principle to a freely falling object. Locally, one can define a coordinate system x^μ such that its equation of motion corresponds to a frame with no acceleration and gravitation, i.e. $d^2x^\mu/d\lambda^2 = 0$ where λ is an evolution parameter which monotically increases along the particle's path as time is now one of the four coordinates.
According to the Equivalence Principle, this equation also holds in the vicinity of any

object, so there is another coordinate system x'^{μ} where we can rewrite this equation as follows :

$$\frac{d^2x^{\mu}}{d\lambda^2} + \Gamma^{\mu}_{\alpha,\beta}\frac{dx^{\alpha}}{d\lambda}\frac{dx^{\beta}}{d\lambda} = 0 \tag{1.3}$$

It corresponds to the geodesic equation, the equation of motion in GR, where $\Gamma^{\mu}_{\alpha,\beta}$ are the *Christoffel symbol*. It is an affine function between the two coordinate systems that correctly describes the effects of parallel-transport in a curved space-time :

$$\Gamma^{\mu}_{\alpha\beta} = \frac{\partial x'^{\mu}}{\partial x^{\nu}}\frac{\partial^2 x^{\nu}}{\partial x'^{\alpha}\partial x'^{\beta}} \tag{1.4}$$

It can also be expressed as a function of the metric and its derivatives. In a cartesian coordinate system, the Christoffel symbol vanishes and the geodesic equation is simply : $d^2x^{\mu}/dt^2 = 0$. The geodesics can be interpreted as the straightest possible paths in a curved geometry.
Although absolute gravitational force has no meaning, the relative gravitational field (sometimes called the tidal field) between two nearby events can be measured by observing the relative acceleration of two freely falling bodies. This relative acceleration is directly related to the curvature of space-time by the geodesic equation deviation. This deviation represents the tendancy of the geodesics to accelerate toward or away from each other due to the curvature of space-time.

Einstein's equation The great advantage of the metric in GR is that it incorporates gravity : instead of considering gravity as an external force with particles moving in a gravitational field, we can include gravity in the metric and describe the motion of free particles in a curved space-time. The notion of geodesics in GR thus implies considering gravity as an aspect of space-time curvature. A second aspect of GR is that one can relate the metric to the matter and energy content of space-time. Such a relation is given by the Einstein's equation which relates the components of the *Einstein tensor* $G_{\mu\nu}$, describing the geometry, to the *energy-momentum tensor* $T_{\mu\nu}$, describing the energy content, through :

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{1.5}$$

where $R_{\mu\nu}$ is the *Ricci tensor* which depends on the metric and its derivatives and R is the *Ricci scalar* which is a contraction of the Ricci tensor ($R = g^{\mu\nu}R_{\mu\nu}$). G is the Newton's constant and $T_{\mu\nu}$ is the energy-momentum tensor which acts as the source of gravitation. So in the GR viewpoint, massive objects distort the space-time curvature and other objects in the vicinity thus follow geodesics in a curved space.

1.1.1.2 Expanding universe

In 1917, just two years after the completion of GR, Einstein introduced the first modern cosmological model and applied the framework of GR to the universe by (correctly) assuming, with little observational guidance at that time, that it is homogeneous and isotropic on large scales [8]. He also assumed (incorrectly) that the universe is static. Finding these two assumptions to be incompatible with the natural solution of his equation, Einstein modified equation 1.5 to include a 'cosmological term', now usually known as the **cosmological constant** and denoted Λ :

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu} \tag{1.6}$$

In effect, he added a new component whose repulsive gravity could balance the attractive gravity of the matter (though he did not describe his modification in these terms). In section 1.1.5, we will see in more detail why we include this cosmological constant now. Before presenting the mathematical tools associated with the application of GR to the whole universe in section 1.1.1.3, let first go back to the concept of expanding universe.

Redshifts in cosmology

The notion of **redshift** z is central in cosmology as most of our observations come from light which is emitted by astrophysical object. Here, we follow the definitions and notations given in [9]. The redshift z of an object is the fractional Doppler shift of its emitted light resulting from radial motion

$$1 + z = \frac{\nu_{\text{R.F.}}}{\nu_{\text{obs}}} = \frac{\lambda_{\text{obs}}}{\lambda_{\text{R.F.}}} \tag{1.7}$$

where $\nu_{\text{R.F.}}$ and $\lambda_{\text{R.F.}}$ are the frequency and wavelength emitted in the rest frame, and ν_{obs} and λ_{obs} are the observed ones. The right panel of figure 1.2 illustrates the shift of spectral lines due to relative motion w.r.t an observer at rest. If a source is moving closer to an observer, its wavelength is smaller compared to a source at rest which translates into a blueshift of the spectral lines. Conversely, if a source is moving away from the observer, its frequency increases and the spectral lines will be redshifted compared to the spectrum of a source at rest. Such effect is known as the *Doppler effect*. In the case of the universe, the spectral lines are redshifted due to the expansion. In fact, a photon's wavelength can be shifted because of three main effects :

- the Doppler redshift due to the relative peculiar velocity between the source and the observer which will contribute to both blue and red Doppler shift
- the **cosmological redshift** which is a pure consequence of the expansion of the universe and thus corresponds to a red shift.
- the gravitational redshift due to the Einstein's effect when there is a difference in magnitude of the gravitational potential between the observer and the photon's source. Einstein took into account this effect to calculate the amount of light deflection by the Sun which was in agreement with observations and considered as the first success of GR. Except in the close vicinity of strong gravitational fields, such as for white dwarf stars, neutron stars and black holes, this effect is negligible and ignored in this thesis.

For low redshifts ($z \ll 1$) which correspond to small v/c (or small distance d in the expanding universe), the following relation holds : $z \simeq v/c$. So a measurement of the amount by which spectral lines of galaxies are redshifted is a direct measure of how fast they are receding from us.

Observational evidence of an expanding universe started emerging in 1912 when Vesto Slipher, director of the Lowell Observatory in Arizona observed the spectral lines of 'spiral nebulae' using a camera and a spectrograph on the 60-cm telescrope at the observatory. He recorded their spectra and was able to measure the spectral shift and infer the recession velocity of those nebulae w.r.t. us (see the box below for an explanation of a spectral

line shift). Slipher [10] found that some nebulae exhibited a shift towards the red part of the wavelength domain with recession velocites up to 1100 km.s^{-1}. It therefore suggested that the universe was far bigger than what the scientific community thought at that time, making the nature of those 'spiral nebulae' a subject under discussion (nowadays, they correspond to objects within our own galaxy or independent galaxies).

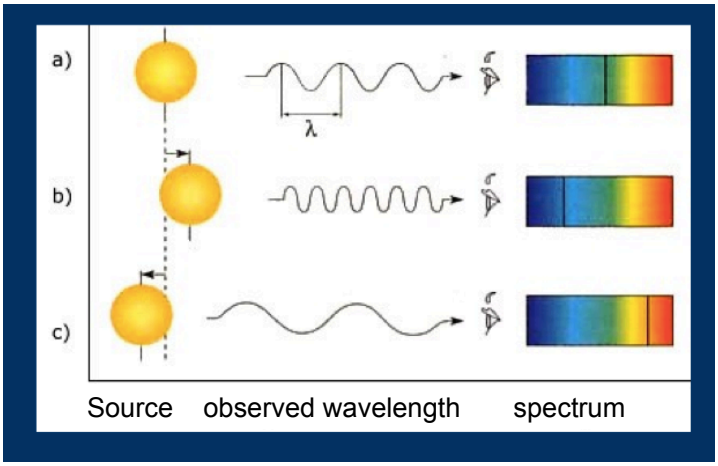


FIGURE 1.2 – Illustration of the Doppler shift of spectral lines due to the relative motion of a source w.r.t. a fixed observer. The shift of spectral lines towards the red part of the spectrum for distant galaxies can be interpreted as a consequence of the expansion of the universe which make galaxies move away from us.

The doubt on the implication of Slipher's redshift measurements remained until 1929 when Edwin Hubble published a seminal article [11] where he simultaneously measured a sample of 25 galaxies. He used the 2.5m-telescope at the mount Wilson in California to distinguish giant bright stars in spiral galaxies whose light curves exhibit a mean magnitude that depends on their pulsation period. Those giant bright stars are called Cepheids and had been studied a few years before by Henrietta Leavitt at the Harvard College Observatory. She determined a relation between their pulsation period and their intrinsic brightness using the variability of Cepheids in the Magellanic Clouds [12]. Edwin Hubble used her relation to measure the distances of galaxies up to ∼15 Mpc as a function of their recession velocities inferred from the redshift of their spectral lines. He showed that when plotting the recession velocity of a galaxy as a function of its distance, he could find a law of proportionality, now known as the *Hubble's law*, meaning that the further galaxies are, the faster they move away from each other. This law of proportionality can be expressed as :

$$v = H_0 D \tag{1.8}$$

where v is the recession velocity, D is the physical distance from the galaxy to the observer and H_0 is called the **Hubble constant** and was measured by Hubble to be 530 km.s^{-1}.Mpc^{-1}. This relation holds for the local universe ($z \ll 1$), we will see a more general expression for the Hubble parameter as a function of time and energy content in section 1.1.1.3. On the theoretical side, in 1922 Alexander Friedmann published a set of equations derived from Einstein's equation showing that the universe might expand [13]. Five years later, Georges Lemaitre interpreted it as the expansion of a homogeneous and isotropic universe with uniform density and estimated the value of the expansion rate to be $H_0 = 570$ km.s^{-1}.Mpc^{-1} [14, 15]. Figure 1.3 shows the calculation of Lemaitre on the left panel while the right panel shows the measurements by Hubble.

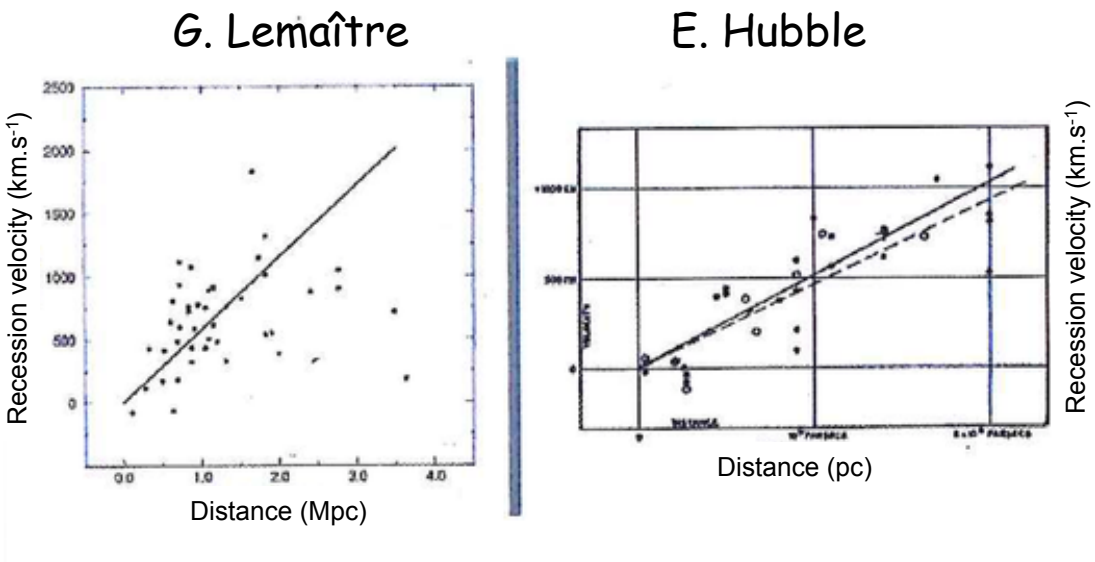


FIGURE 1.3 – Hubble diagram, i.e. recession velocities as a function of distance, as calculated by Georges Lemaitre in the left panel and as measured by Edwin Hubble in the right panel. Both found a consistent coefficient of proportionality which is now known as the Hubble constant.

Georges Lemaitre also first proposed the idea of a primeval atom to highlight the fact that if the universe is expanding, it can have a beginning and so a history. In fact, the expansion of the universe which is illustrated in the Hubble diagrams of figure 1.3 represents the first pillar of what will be called the *hot Big Bang model*. An usual image to describe the expansion of the universe is a cooking raisin cake, as illustrated in figure 1.4 where the cake inflates and each raisin can see its neighbours moving away from it, all the more quickly as they are far, but their own size doesn't change.

A series of H_0 measurements has been performed after Lemaitre and Hubble. In the three decades that followed, published values of the Hubble constant varied by about a factor of two between ∼100 and 50 km.s^{-1}Mpc^{-1} as summarized in [e.g. 16]. In order to represent the spread in the measurements of the Hubble constant, a useful notation is :

$$H_0 = 100h\,\mathrm{km.s^{-1}.Mpc^{-1}} \tag{1.9}$$

It means that in our today universe, two galaxies that are distant by 1 Mpc move away from each other with a velocity of $100\,h$km.s^{-1}. Depending on the quantity of interest, the inverse of the Hubble constant, H_0^{-1} gives access to the Hubble time t_H and times the speed of light to the Hubble distance D_H :

$$t_H = H_0^{-1} = 9.78\,h^{-1} \times 10^9\mathrm{yr} \tag{1.10}$$

$$D_H = \frac{c}{H_0} = 3000\,h^{-1}\mathrm{Mpc} = 9.25 \times 10^{25}\,h^{-1}\mathrm{m} \tag{1.11}$$

Taking $H_0 = 530\,\mathrm{km.s^{-1}}$ as measured by Edwin Hubble, it gives $H_0^{-1} = 2 \times 10^9$yr which corresponds to the age of the universe assuming a constant recession velocity and a matter-dominated only universe at present epoch . Already in the 1930's, this estimation of the age of the universe was in disagreement with the geophysics estimate of the age of the Solar System to be at $4-5\times10^9$yr. It was an argument in favor of steady state theories [17, 18] where the universe, although in expansion, remained the same in time with no beginning

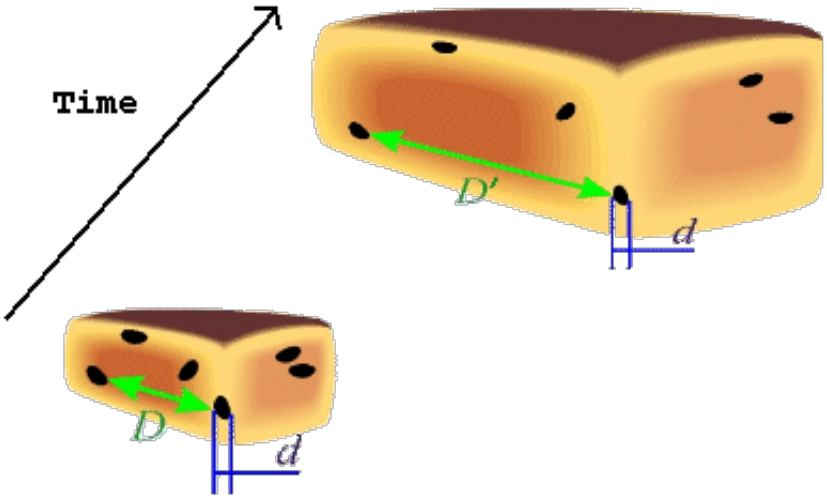


FIGURE 1.4 – Illustration of the expansion of the universe using the exemple of the cooking raisin cake. As time goes, the distances between raisins (galaxies) increase but not the size of the raisin as it is bound by electric (gravitational) interactions between its components.

and no end. This model remained popular until the discovery of the black-body radiation of photons that were emitted in the early universe, 380000 years after what represents the $t = 0$, the beginning of the universe called the Big Bang.

1.1.1.3 Dynamics of the universe

Cosmological principles To describe the dynamics of the universe, we need to make some assumptions on its geometry which are now confirmed by independent observations of the universe at different epochs and scales. Such assumptions constitute the Cosmological Principle which states that no observer occupies a preferred position in the universe. This principle thus implies that the universe is **homogeneous** and **isotropic** (no preferred direction in the sky or orientation on the sky) on scales of $> 100\,\mathrm{Mpc}$, i.e. scales much larger than the typical distance between galaxies. These characteristics define the *background universe*.

Friedmann-Lemaitre-Robertson-Walker metric From a mathematical viewpoint, the isotropy and homogeneity of the universe imply that only two parameters fully determine its geometry : a factor to describe its expansion which can only be a function of time because of homogeneity and a global curvature k which must be the same everywhere because of isotropy. To describe an expanding universe, we could modify the metric by multiplying the spatial terms with a time-dependent factor $R(t)$. In spherical coordinates, it gives :

$$ds^2 = c^2dt^2 - R^2(t)(dr^2 + r^2d\theta^2 + r^2\sin\theta^2 d\phi^2) \tag{1.12}$$

To describe the size of the universe, it is more convenient to define the **scale factor** a as :

$$a(t) = \frac{R(t)}{R_0} \tag{1.13}$$

where the subscript 0 refers to the present epoch and by definition, $a(t_0) = 1$. The scale factor relates the physical distances D_{phys} to the coordinate comoving distances by $D_{\mathrm{phys}} =$

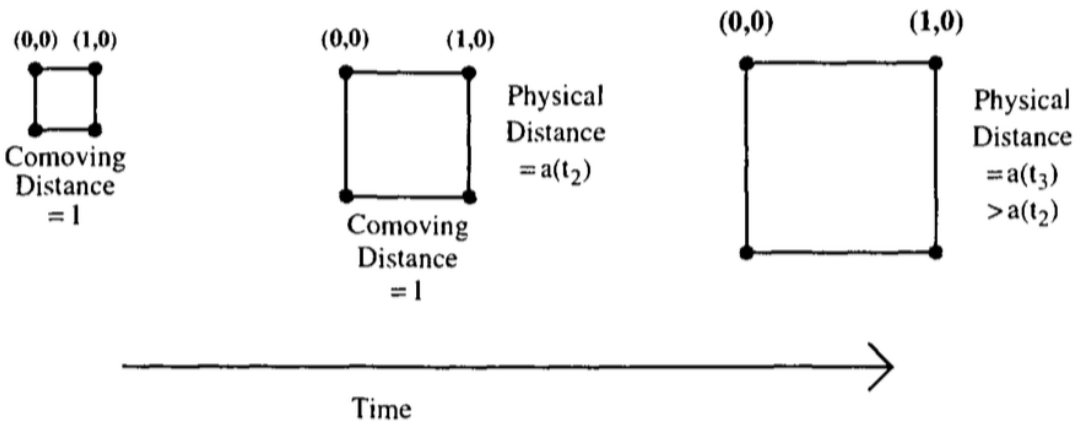


FIGURE 1.5 – We define a comoving distance whose value remains constant as the universe expands. From [4].

$a(t)D_C$ where D_C is the comoving distance defined by equation 1.23. Figure 1.5 illustrates the expansion of the universe as a grid which expands uniformly as time evolves. Points on the grids maintain their coordinates so the *comoving distance* between two points remains constant. The scale factor is also directly related to the cosmological redshift introduced in the previous section by :

$$1 + z = \frac{a(t_0)}{a(t)} \tag{1.14}$$

where at t_0, $a(t_0) = 1$ and $z = 0$ corresponds to the present epoch.
In fact, the most general metric that describes an homogeneous, isotropic and expanding universe is :

$$ds^2 = c^2dt^2 - a(t)\left(\frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2\sin^2\theta d\phi^2\right) \tag{1.15}$$

where the constant k determines the curvature of the universe : $k = 0$ for a flat universe, $k = +1$ if spherical or $k = -1$ if hyperbolic. Equation 1.15 corresponds to the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric.

Friedmann equations The Friedmann equations are the key equations that describe the dynamics of the universe as a whole. Proving these equations would take us a long way outside the scope of this manuscript into what is usually graduate-level physics, so we refer the reader to the mentioned textbooks for their derivation from Einstein's equations by imposing local energy conservation. The simplest form of $T_{\mu\nu}$ is the one that describes a perfect-fluid with no viscosity whose velocity vector coordinates are denoted u^μ. Under this assumption, the energy tensor of the universe becomes

$$T^{\mu,\nu} = \rho[(w+1)u^\mu u^\nu - wg^{\mu\nu}] \tag{1.16}$$

where w is the equation of state parameter that relates the pressure p to the density ρ by $p = w\rho c^2$.
The time component ($\mu = \nu = 0$) of the Einstein's equations (1.5) describes how fast the universe is expanding :

$$\frac{1}{a^2}\left(\frac{da}{dt}\right)^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - kc^2 \tag{1.17}$$

and by combining the time and the space components, we find an equation for the deceleration of the expansion of the universe :

$$\frac{\ddot{a}}{a} = \frac{-4\pi G}{3c^2}(\rho + 3p) \tag{1.18}$$

For a given fluid whose equation of state is $p = w\rho c^2$, combining equations 1.17 and 1.18 gives :

$$\frac{\dot{\rho}}{\rho} = -3(w+1)\frac{\dot{a}}{a} \implies \rho(t) = \rho(t_0)a(t)^{-3(w+1)} \tag{1.19}$$

It implies :

$\rho_m(a) = \rho_m(a_0)a^{-3}$	$\lvert p\rvert \ll \rho c^2$	non-relativistic matter : baryons + cold dark matter
$\rho_r(a) = \rho_r(a_0)a^{-4}$	$p = \rho c^2/3$	radiation / relativistic matter (photons + massless neutrinos)
$\rho_\Lambda(a) = \rho_\Lambda(a_0)$	$p = -\rho c^2$	vacuum / cosmological constant

The fact the vacuum energy density is constant is due to its value being independent of the number of particles present. The cosmological constant is then interpreted as a fluid of constant density with negative pressure.

We can also define the total density parameter $\Omega_{tot} = \Omega_m + \Omega_r + \Omega_\Lambda = 1 - \Omega_k$ and the critical density $\rho_{\text{crit}} = 3H^2/8\pi G$, such that $\Omega_{m,r} = \rho_{m,r}/\rho_{\text{crit}}$. The density parameters for each component can be written as :

$$\Omega_{m,r,\Lambda} = \frac{\rho_{m,r,\Lambda}}{\rho_{\text{crit}}} \tag{1.20}$$

$$\Omega_k = \frac{-kc^2}{R^2H^2} \tag{1.21}$$

where H is the **expansion rate** (also called the Hubble parameter in the previous section) and which is defined by $H(t) = \dot{a}/a$. Including the time-evolution of the density parameters, one can therefore rewrite the 1^{st} Friedmann equation (1.17) by :

$$H^2(t) = H_0^2[\Omega_{m,0}a(t)^{-3} + \Omega_{r,0}a(t)^{-4} + \Omega_{\Lambda,0} + \Omega_{k,0}a(t)^{-2}] \tag{1.22}$$

As we said in the introduction, due to the the change in temperature and pression conditions as the universe expands, the densities of the different components evolved with time and changed their relative ratios such that the universe successively underwent eras of radiation, matter and dark energy domination. This is illustrated by figure 1.6 which represents the redshift evolution of each density component and their relative contribution to the total density for a flat universe ($k = 0$).

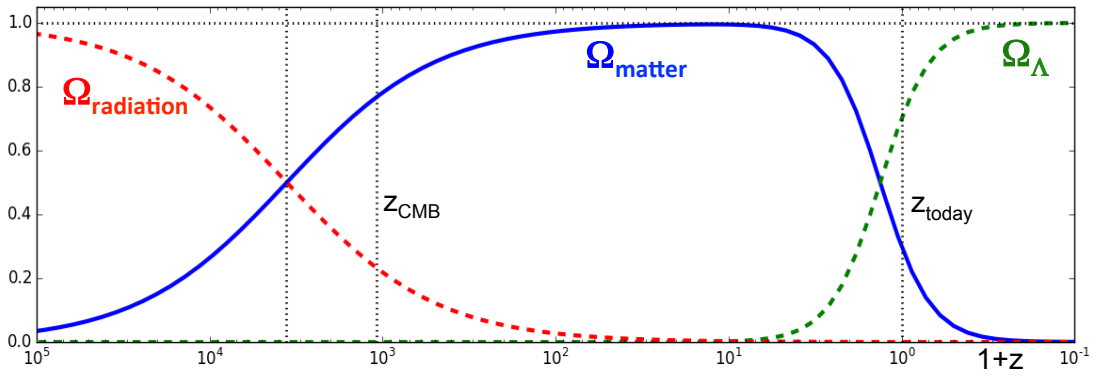


FIGURE 1.6 – Radiation density (red), Matter density (blue) and cosmological constant density (green) as a function of redshift for a flat universe $\Omega_{tot} = \Omega_m + \Omega_r + \Omega_\Lambda = 1$. The redshift at CMB (just after the matter-radiation equality) and the today redshift are also represented.

Distances in cosmology

LOS comoving distance D_C Defined as the distance between a distant emitter at scale factor a and us,

$$D_C = c \int_0^z \frac{dz'}{H(z')} \tag{1.23}$$

where $H(z)$ depends on the energy content of the universe and is defined in equation 1.22. The radius of the *observable* universe is obtained by integrating to infinity. For the currently-accepted values of the density parameters [19], it comes out about $3.5c/H_0$ and the volume enclosed in this radius is referred to as the Hubble volume.

Transverse comoving distance D_M Defined as the ratio between the comoving distance between two events at the same redshift and their angular separation $\Delta\theta$. It is related to the **angular diameter distance** D_A through

$$D_A = \frac{D_M}{1+z} \tag{1.24}$$

1.1.2 The hot Big Bang model

The *hot Big Bang model* (HBB) postulates that the Universe started 13.8 billion years ago when it was much hotter and denser than today. It has been developed in the 1950s by Georges Gamow whose doctoral advisor was Alexander Friedmann. He applied Friedmann's and Lemaitre's non-static solutions to describe a universe of uniform density and constant spatial curvature. He also extended Lemaitre's idea of a primeval atom by assuming that the early universe was dominated by radiation rather than matter. He and his students introduced the idea of primordial nucleosynthesis during which the light elements were formed and they also predicted the existence of a relic thermal (blackbody) spectrum of photons, the cosmic microwave background (CMB) [20, 21, 22]. So by 1970's, the HBB

theory is supported by three major observations :

— the Hubble law which is still the most direct evidence of the expansion of the universe

— the abundances of light elements that were formed during the primordial nucleosynthesis, also called Big Bang nucleosynthesis (BBN)

— the black-body radiation of the CMB as a relic of the thermal photons emitted during the early universe

In this section, we will just provide the most relevant results related to the BBN and the CMB.

1.1.2.1 Big Bang Nucleosynthesis

In the early universe, temperatures were much hotter than today and the thermodynamics was therefore very different than today where photons and neutrinos very rarely scatter on matter. Knowing the conditions of the early universe and the relevant nuclear cross-sections, we can calculate the expected primordial abundances of all elements as a function of the density of protons and neutrons, and hence on the density of baryons compared to photons, at the time of nucleosynthesis.
The key quantity to understand what happens in terms of nuclear reations is the comparison between the numbers of reactions per particle per unit time and the Hubble rate whose inverse gives the characteristic time $t_H = H^{-1}$ for temperature and density changes due to the expansion. Given this consideration, the thermal history of the early universe can be summarized into three steps which are illustrated in figure 1.7 :

(i) **Thermal equilibrium** occurs until the neutron-proton *freeze-out* for $T < 1\,\mathrm{MeV}$ and $t < 1\,\mathrm{s}$: all the relativistic components of the universe are in thermal equilibrium (all reactions proceed with the same rate in forward and backward directions) and the universe is dominated by radiation. The universe is too hot and too dense to have neutral atoms or even bound nuclei. More precisely two equilibriums are reached : kinetic equilibrium due to elastic scattering with energy exchanges between particles and chemical equilibrium due to inelastic collisions where the number of particles is changed :

$$p + e^- \leftrightarrow n + \nu_e \tag{1.25}$$
$$p + \bar{\nu_e} \leftrightarrow n + e^+ \tag{1.26}$$
$$\gamma + e^- \leftrightarrow \gamma + + e^- \tag{1.27}$$
$$p + e^- \leftrightarrow p + e^- + \gamma \tag{1.28}$$
$$\gamma + \gamma \leftrightarrow e^+ + e^- \leftrightarrow \nu_e + \bar{\nu_e} \tag{1.29}$$

At $T \sim 1\,\mathrm{MeV}$, the reaction time scale for $p + e^- \leftrightarrow n + \nu_e$ became longer than t_H and after this point, the neutron-proton reactions cease and the neutron-to-proton ratio $\mathrm{n}_n/\mathrm{n}_p$ is frozen out at the equilibrium value that it had at that time. It is followed by the neutrino freeze-out shorlty after.

(ii) **Neutrons decay** for $800\,\mathrm{keV} > T > 60\,\mathrm{keV}$: Neutrons are not stable and instead decay into protons $n \rightarrow p + e^- + \bar{\nu_e}$ with a lifetime of ∼880 s. Then, the nuclear reactions during the next ∼1000 s determined the relative abundances.

(iii) **Primordial nucleosynthesis** started when $T \sim 60\,\mathrm{keV}$ until the freeze-out of nuclear reactions at $T \sim 30\,\mathrm{keV}$ (so between ∼200-1000 s) resulting in the production of substantial amounts of deuterium D ($D/H \simeq$ few 10^{-5}), helium ^3He and ^4He, and lithium Li. Primordial nucleosynthesis could not start earlier because for $T \sim$ MeV

because of high temperatures and energetic photons which did not allow to the formation rate of deuterium to be greater thant the destruction rate by photo-dissociation. The formation of light nuclei stopped at ^7Li because of the absence of stable elements at $A = 5, 8$ and then because the temperature and density conditions were no longer satisfied. The synthesis of heavier elements started up again once stars were formed initiating the stellar nucleosynthesis.

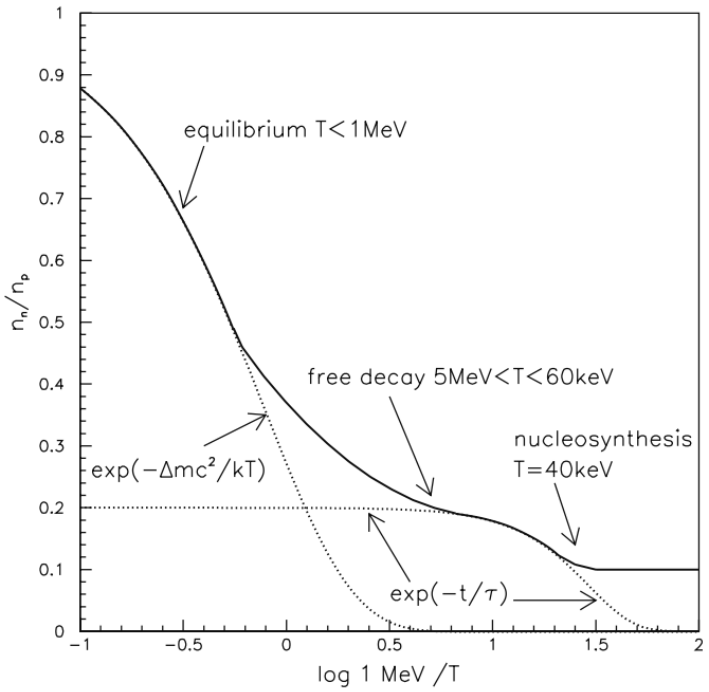


FIGURE 1.7 – The neutron-proton ratio as a function of time for temperatures between 10 keV and 10 MeV. The three main steps of BBN are represented : thermal equilibrium, free decay and synthesis of the first light nuclei until $T \sim 30\,\text{keV}$. From [5].

All theoretical predictions of primordial abundances agree well with the measurements, except for ^7Li where the observations find a significantly lower abundance. However, this seems to be due to astrophysical effects related to the stellar nucleosynthesis that can affect the measurements of primordial abundances.

1.1.2.2 Cosmic Microwave Background

Just after the synthesis of the first light nuclei, the universe was still dominated by radiation and photons are still energetic enough to ionize the recently-formed atoms. In the early universe, due to high temperatures, baryonic matter existed in form of neutrons and charged protons and electrons. During the BBN, the neutrons and protons formed charged atomic nuclei, mostly hydrogen and helium ions. But atoms could not be formed due to absorbed or scattered thermal photons whose mean energy was above the ionization level of atoms. Until recombination, the baryonic matter was therefore strongly coupled with the photons by Thomson scattering with free electrons and in general, all charged particles were exchanging energy continously in frequent collisions, so that the whole formed an ionized plasma. This is called the **baryon-photon plasma**.
In this plasma, density perturbations were damped compared to CDM perturbations as the

baryon-photon pressure prevents the growth due to gravitational instability. This counteracting effects gave rise to acoustic waves that propagated away from the density peaks in the plasma and that are called **baryon acoustic oscillations, BAO**. We will go back to the imprint of BAO in the present day matter clustering in section 1.2.2.1.
As the universe became cold enough that thermal photons were no longer energetic enough to ionize particle composites, free particles started forming atoms. It marks an epoch called **recombination** associated with the *baryon-photon decoupling* which allows the mean free path of photons to become larger than the Hubble horizon so light could cross the universe without scattering. From the time of decoupling until today, these photons were redshifted and formed the cosmic microwave background (CMB). Because photons were in kinetic equilibrium before recombination and given that they are massless, the thermal character of their spectrum is maintened by the expansion. The CMB is therefore the redshifted black body photon radiation from when the universe was last opaque. Because of the finite speed of light, the CMB must appear to any observer as originating from a spherical receding surface, with the observer at the center, called the *surface of last scattering*.

The CMB was discovered in 1964 by Arno Penzias and Robert Wilson [23, 24] at the Bell Laboratories in New Jersey while they were calibrating a microwave antenna for use in telecommunications and astronomy. They found a surprising and unexplained constant noise source distributed isotropically across the sky. At the same time, in Princeton a group led by Robert Dicke [25] had just predicted an isotropic CMB from the Big Bang theory [3]. After the discovery, the three major space missions that were dedicated to CMB observations are the NASA's satellite COsmic Background Explorer [26, *COBE*,] from 1989 to 2007, then the *WMAP* satellite [27] from 2001 to 2007 and the ESA's *Planck* mission from 2009 to 2013 whose latest cosmological results have been published in 2015 and are described in [19]. The energy content of the universe shown in figure 1.1 is based on the *Planck* 2015 results and corresponds to the flat ΛCDM model used in this manuscript.

Figure 1.8 shows the intensity of the CMB radiation measured by *COBE* as a function of frequency. It is the most perfect black body spectrum ever known with a mean temperature $T_{\rm obs} = 2,725K$ and temperature anisotropies at the level $\Delta T/T \sim 10^{-5}$. Since photon energy is related to frequency ν by $E = h\nu$ where h is the Planck constant, the energy density has decreased by a factor of $(1+z)^4$ since the photons were emitted (because volumes have decreased by a factor of $(1+z)^3$ and wavelengths by $(1+z)$). The Stefan-Boltzmann law states that the energy density is proportional to T^4 so at $z_{\rm CMB} \sim 1090$, $T_{\rm em} = (1+z)T_{\rm obs} \sim 3000K$.

To describe the temperature fluctuations across the sky, the signal must be decomposed onto a basis which is appropriate for the spherical surface of the sky. Such a basis is called the spherical harmonics Y_{lm} and the temperature anisotropies can thus be written :

$$\frac{\Delta T}{T}(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm}Y_{lm}(\theta,\phi) \tag{1.30}$$

where θ,ϕ the angles for spherical coordinates and the a_{lm} are the amplitudes of each l,m mode which are very close to be Gaussian. In addition, $\langle \Delta T/T \rangle = 0$ so that $\langle a_{lm} \rangle = 0$, thus all the information is encoded in the variance, i.e. the correlation between two points in the sky, for all the points. This statistical tool is called the *two-point correlation function*

3. Gamow's prediction was implicit in calculations published in 1948

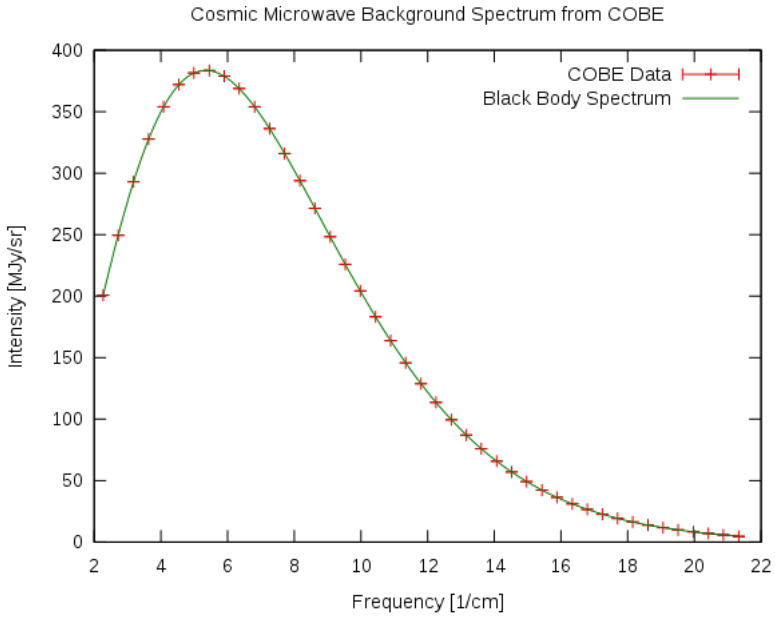


FIGURE 1.8 – The intensity of the CMB radiation as measured by *COBE* as a function of frequency compared with a theoretical black body spectrum. From [28]

of temperatures fluctuations and it is defined by

$$C(\eta_1, \eta_2) = \langle \frac{\Delta T}{T}(\eta_1) \frac{\Delta T}{T}(\eta_2) \rangle \tag{1.31}$$

where $\eta = (\theta, \phi)$. One can also parametrize the direction (η_1, η_2) by one value, its dot product $\cos\alpha$ so that :

$$C(\eta_1, \eta_2) = C(\cos\alpha) = \sum_l \frac{2l'+1}{4\pi} C_{l'} P_{l'}(\cos\alpha) \tag{1.32}$$

where $P_{l'}(\cos\alpha)$ can be decomposed into spherical Legendre polynomials. One can then show that $C_l = \langle |a_{lm}|^2 \rangle_m$. Only the $2l+1$ independent m modes can be measured at any fixed l, which limits the precision of the measurements of C_l. This limit is known as **cosmic variance** and comes out as

$$\Delta C_l = \sqrt{\frac{2}{2l+1}} C_l \tag{1.33}$$

It corresponds to the best possible measurement in the absence of any instrumental noise or astrophysical systematics.
Figure 1.9 represents the amplitude of temperature fluctuations $D_l^{TT} = l(l+1)C_l/2\pi$ as a function of angular scales l (or θ[rad] $\sim 180/l$) as measured by [19].

An important scale is the angular size of the maximum distance particles can travel until the time of decoupling. It corresponds to 2° in the sky and modes further apart than 2° could not have been in causal contact and are called *super-horizon* modes.
In what follows, we will describe the three main features of the CMB spectrum.

— (1) **Sachs-Wolfe plateau** : On scales much larger than the size of the horizon at decoupling, the only causal contact could have been before the radiation era, i.e. during a period called *inflation* and that will be described in the next section. Inflation must predict *scale-invariant* fluctuations to be consistent with the constant level of the temperature fluctuations on large scales (low l) that are seen in the CMB. This plateau is related to the Sachs Wolfe (SW) effect which is due to the gravitational redshift of photons at last scattering that have to escape from the potential wells where baryons are. This ordinary SW effect is following by an early integrated SW effect and then a late ISW. The ISW is due to the change in energy of the CMB photons that travel from last scattering to the present by crossing the evolving metric potentials : the early ISW produces an enhancement of the acoustic peak at $l \sim 200$ when there is still radiation in the universe and a late ISW that is responsible for the raise at $l < 30$.

— (2) **Acoustic peaks** : Before recombination, the motions of baryons and photons are strongly coupled, whereas cold dark matter decoupled earlier and started collapsing. Due to gravitational interaction, the baryon-photon plasma falls out of under-dense regions and into over-dense regions where the gas compressed towards the center of the over-dense regions. But photon pressure outwards resisted the inward flow, which sets up oscillations. These oscillations propagate during the first 380,000 years until decoupling when oscillations were frozen, which sets the sound horizon size at decoupling r_s which gives the approximate position of the first acoustic peak. The sound horizon at decoupling $z_{\rm dec} = 1090$ is defined by :

$$r_s(z_{\rm dec}) = \int_{z_{\rm dec}}^{\infty} \frac{c_s(z)}{H(z)} dz = (97.34 \pm 0.33)\, h^{-1}{\rm Mpc} \tag{1.34}$$

where $c_s \simeq c/\sqrt{3}$ is the speed of sound which depends on the ratio of the radiation pressure to the energy density of the baryon-photon fluid, determined by the baryon-to-photon ratio which is proportional to $\Omega_b h^2$. $H(z)$ is the expansion rate whose behavior at $z > z_{\rm dec}$ depends on the ratio of the matter density to radiation density, where for a fixed radiation sector ratio, it is proportional to $\Omega_m h^2$. Both ratios are well measured by the relative heights of the acoustic peaks in the CMB anisotropy power spectrum.

— (3) **Silk damping** : The higher acoustic peaks are suppressed because of photons diffusion on the last scattering surface. In fact, the transition from an opaque to a transparent universe was not instantaneous and during the transition, photons were diffusing away which smoothed out the structure on the smallest-scales (reducing the amplitude of their fluctuations). This effect is known as diffusion damping or Silk damping [29].

Summary

The hot Big Bang model is a very satisfactory picture of the universe for $T \leq 100\,{\rm MeV}$ and $t \geq 10^{-4}\,{\rm s}$, which assumes general relativity as the underlying theory of gravitation, the cosmological principle (homogeneity and isotropy) and the known particle physics. We have seen that it is consistent with many observations : the Hubble's law that reveals the expansion of the universe, the abundance of light elements that are formed during the Big Bang nucleosynthesis just few minutes after the Big Bang, the existence of cosmic microwave background which is testament to the recombination epoch when the universe was 380,000 years old and offers a wealth of information on the initial conditions for structure formation. In addition, we will see in section 1.1.4 that the growth of perturbations are

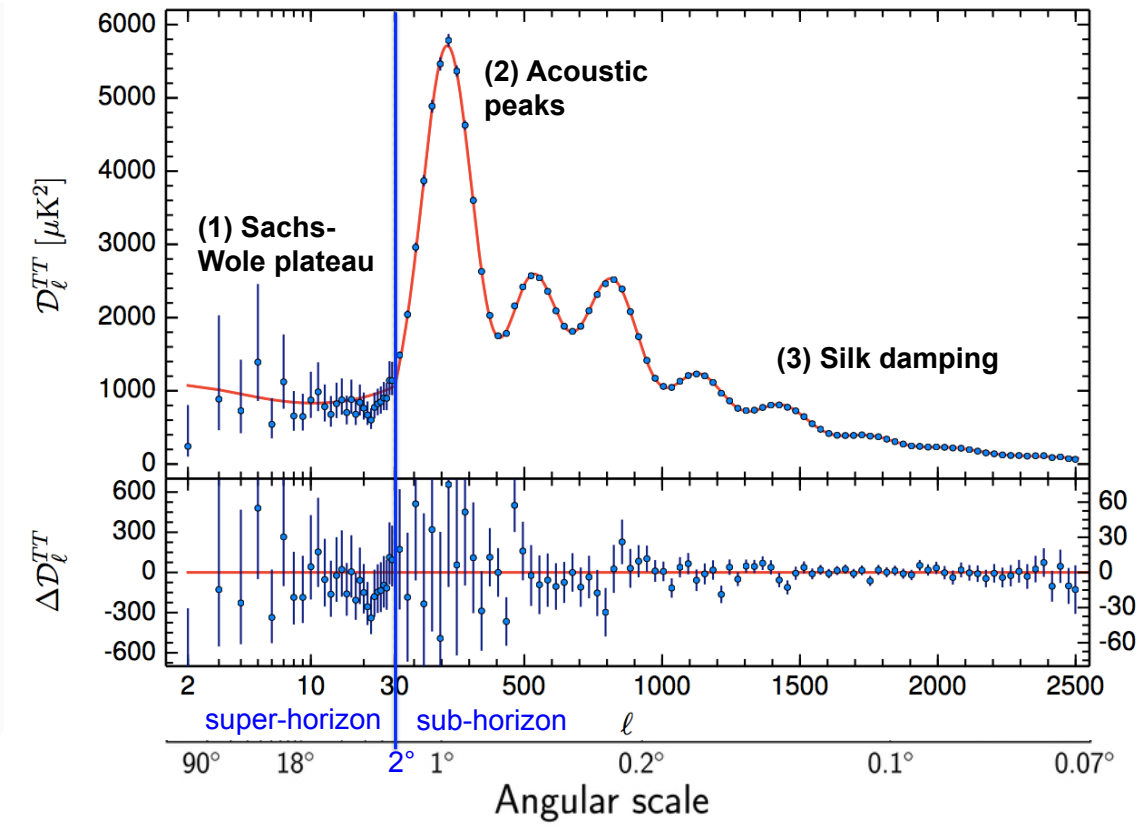


FIGURE 1.9 – Adapted from [19].

well-explained by gravitational instability in an expanding universe with an approximate scale-invariant initial power spectrum. In the hot Big Band model, the universe emerged from a hot and dense phase in thermal equilibrium but the origin of these initial conditions are not explained by the model, i.e. we don't know the origin of the primordial fluctuations. There are also two other major unanswered questions :

— *Horizon problem* : we already said that the angular size of the horizon at recombination corresponds to 2° in the sky, so regions further apart than are not causally-connected in the sense that we do not expect to find a similar behavior between these regions. So the uniformity of the CMB across the full sky cannot be explained since no causal mechanism could have made the temperature uniform on all scales, including super-horizon ones.

— *Flatness problem* : CMB observations of the temperature fluctuations alone favor a flat universe with $\Omega_k = -0.052^{0.049}_{0.055}$ [19]. It implies that in the past $\Omega_m = 1$ at the level of 10^{-60}, meaning that the value of Ω_m must be adjusted very precisely in order to fit with CMB observations and tiny deviations from this value would have led to very different universes. When some conditions appear to be fine-tuned to specific values, we call this problem a *fine-tuning* of the initial conditions.

1.1.3 Primordial universe

In order to avoid an extreme fine-tuning of the curvature at early times and to understand how scales which seem uncorrelated today are observed to have almost identical temperatures in the early universe, we need a mechanism that acts as a very rapid expansion of the universe before the radiation era. It corresponds to a period of **inflation** where $a(t) \sim t^p$ with $p > 1$, which implies $\ddot{a} > 0$ (acceleration of the expansion). We will see that such a mechanism can also explain the origin of fluctuations, but at the time of writing this thesis it is still not sure whether inflation is the right mechanism.

1.1.3.1 Inflation

The idea behind inflation that was first proposed by [30] is to assume that there was an epoch when vacuum energy was the dominant component of the energy density of the universe so that the scale factor grew exponentially. In this approach, the universe is filled with a *scalar* field which, contrary to *vector* fields such as the gravitational force or any of the four fondamental force, has no direction. Scalar fields are characterized by an intensity at every point in space and the only scalar field we have detected is the Higgs boson with the Large Hadron Collider (LHC) in 2012. The scalar field associated with inflation has an intensity that evolves only slowly in time in order to explain the horizon and flat problems. The simplest inflationary model introduces a single scalar field, but one can imagine working with multiple fields. Let ϕ denote the value of the scalar field whose associated particle is called the **inflaton** and ϕ is thus called the **inflaton field**. Its associated potential is denoted $V(\phi)$, but we do not know its shape so all the solutions for ϕ will depend on the assumptions made on $V(\phi)$. Here, we will just review the most important aspects of inflation in order to understand the origin of primordial fluctuations and the shape of the initial power spectrum. The theoretical framework of inflation relies on a quantum description of the inflaton field which is beyond the scope of this manuscript, we refer the reader to [2, 31] for a detailed discussion.

As for all quantum fields, the inflaton field is subject to quantum fluctuations that can be decomposed into Fourier modes with amplitude $\psi_k(t)$. Their expressions are obtai-

ned from the equation of motion for a quantum field (the Klein-Gordon equation) in an expanding universe :

$$\ddot{\phi} + 3H\dot{\phi} - \nabla^2\phi + V'(\phi) = 0 \tag{1.35}$$

where H is the expansion rate and ∇ is w.r.t spatial coordinates but since we assume that ϕ is homogeneous, $\nabla^2\phi \simeq 0$. The term $3H\dot{\phi}$ corresponds to a friction term due to expansion.
Equation 1.35 has two different regimes :

- *Slow-roll* regime when $\ddot{\phi}$ is negligible : as long as the kinetic energy ($\propto \dot{\phi}^2$) is small compared to its potential energy ($V(\phi)$, inflation occurs. This slow evolution of the scalar field to the minimum of the scalar potential is a key feature of inflation and was first proposed in [32, 33]
- *Coherent oscillations* when $V'' \geq H^2$ and ϕ evolves rapidly on the expansion time scale, the inflaton will oscillate around a minimum with oscillations that are damped by the $H\dot{\phi}$ term. The inflaton is expected to decay into matter and radiation but the exact mechanism is not known as the underlying physics of the inflaton field is not known. The end result of inflation is that the universe is left with approximately the same energy density as when it started, but in the form of radiation and particles.

For a de-Sitter metric (i.e. for a flat universe with a constant expansion rate), the solutions [4] depend on the factor k/aH where k is the wavenumber meaning that there are two regimes for the evolution of quantum fluctuations. For $k/aH \ll 1$, i.e. for comoving k-modes which exit the Hubble radius $\propto H^{-1}$, the quantum nature of the fluctuations has disappeared and the field ϕ can be seen as a classic stochastic field. If we write the inflaton field : $\phi(t, \mathbf{x}) = \bar{\phi}(t) + \chi(t, \mathbf{x})$, one can show that the perturbations $\chi(t, \mathbf{x})$ around the background homogeneous solution $\bar{\phi}(t)$ generate the primordial fluctuations we observe in the CMB. After the inflationary phase, the modes re-enter the Hubble radius (i.e. $k/aH \gg 1$) and their initial amplitudes at that entry are thus the initial conditions for the HBB model.

1.1.3.2 Statistical properties of cosmic fields

We have seen that for inflationary models, the stochastic properties of the cosmic fields come from quantum fluctuations of the inflaton field. Therefore, in order to describe the distribution of large-scale structures, one needs to adopt a statistical approach and to resort to a probabilistic description. In this section, we will just provide the most important definitions and relevant properties for cosmology.
Let $f(\mathbf{x})$ denote a cosmic scalar random field, meaning its value in each point can be treated as a stochastic variable and $f(\mathbf{x})$ can represent either the cosmological density contrast $\delta_\mathbf{x}$, the gravitational potential $\phi_\mathbf{x}$, the velocity divergence field $\theta_\mathbf{x}$ or any other field of interest. Moreover, inflationary models predict that the initial density perturbations arise from a large number of independent quantum fluctuations.. The central limit theorem implies that a density distribution is asymptotically Gaussian in the limit where the density results from the average of many independent processes. So the initial density perturbations are very close to be Gaussian-distributed, which is in agreeement with observational constraints [19]. It means that cosmic fields are well represented by **Gaussian random fields** (GRF) whose probability distribution function (pdf) is :

$$p(f|\mu, C) = \frac{1}{\sqrt{|2\pi C|}} \exp(-\frac{1}{2}(f-\mu)^T C^{-1}(f-\mu)) \tag{1.36}$$

4. Expressions can be found in [3]

This GRF has mean value $\int_x xp(x|\mu, C)dx = \mu$ and variance $\int_x (x-\mu)(x-\mu)^T p(x|\mu, C)dx = C$. C is assumed to be positive-definite in order to be invertible.
The n-dimensional vector $f = (f_n)$ is a GRF if its joint multivariate pdf can be written as a multi-variate Gaussian following equation 1.36 :

$$P(f|\mu, C) = \frac{1}{\sqrt{|2\pi C|}} \exp(-\frac{1}{2}\sum_{i,j}^{n}(f_i - \mu_i)^T C_{ij}^{-1}(f_j - \mu_j)) \tag{1.37}$$

The cosmological principle implies that the cosmic random field is *statistically homogeneous* (meaning that all the joint pdf remain the same under translation of the coordinates in space) and *statistically isotropic* (invariant under spatial rotations).

In what follows, we will consider the **density perturbation field** δ (also called the density contrast) as defined by :

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}} \tag{1.38}$$

To confront theory that predicts the initial conditions of the statistical distribution of the density contrast to data that provide one measurement of this density field, we need to imagine an *ensemble of universes*. We assume that the observed density field is just one of an ensemble of an infinite number of possible realizations that could have resulted from a random process. Knowing the random process means knowing the probability distribution of ρ that produced the observed field.
However, we have no evidence that the ensemble exists, and we are only able to observe one realization. So the second assumption we need to make is called **ergodicity** which assumes an equivalence between **volume average** and *ensemble average*. The ensemble average is also called the expectation value. If we measure the variance by averaging over a sufficiently large volume, the results would be expected to approach the true ensemble variance and the averaging operator $\langle x \rangle$ is often used without being specific about which kind of average is considered.
Fields that satisfy this property : 'volume average' $\leftrightarrow$ 'ensemble average' are called ergodic. In cosmology, it is regarded as a common sense axiom.

The density perturbation field can thus be written :

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \langle\rho\rangle}{\langle\rho\rangle} \tag{1.39}$$

with from statistical homogeneity $\langle\rho(\mathbf{x})\rangle = \langle\rho\rangle$ which yields $\langle\delta\rangle = 0$. Because a GRF field has no information contained in the phases (i.e. they are all uniformly randomly distributed and because the mean over-density $\langle\delta\rangle = 0$, all the statistical information about the density perturbation field is contained in the variance of its amplitudes, i.e. the **power spectrum** or **two-point correlation function** completely characterizes the density fluctuations.

Fourier convention

There are two conventions in the litterature for the Fourier (k) and configuration (r or s) spaces. In this manuscript I use the following :

$$A(\mathbf{k},t) = \int d^3\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} A(\mathbf{x},t) \tag{1.40}$$

where $A(\mathbf{k},t)$ is called the Fourier transform of $A(\mathbf{x},t)$, $A(\mathbf{k},t)$ is the amplitude of the Fourier mode and $e^{-i\mathbf{k}\cdot\mathbf{x}}$ its phase. For reference, this is the same convention as the one used in [34] which is different from the one used in [3].

Using the previous convention, the power spectrum is defined as and

$$\langle \delta(\mathbf{k})\delta(\mathbf{k}')\rangle = (2\pi)^3 \delta_{\rm D}(\mathbf{k}-\mathbf{k}')P(k) \tag{1.41}$$

where $\delta_{\rm D}$ is the Dirac distribution Its analogue in configuration space is obtained by Fourier transformation and is called the two-point correlation function. It is defined as the joint ensemble average of the density at two different locations :

$$\xi(r) = \frac{1}{(2\pi)^3}\int d^3\mathbf{k} P(k) e^{i\mathbf{k}\cdot\mathbf{r}} = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle \tag{1.42}$$

which depends only on the magnitude of $\mathbf{r}$ due to statistical homogeneity and isotropy. The physical interpretation of $\xi(s)$ is that it measures the excess over random probability that two objects in volume elements dV_1 and dV_2 are separated by the distance $r = |\mathbf{x}_1 - \mathbf{x}_2|$,

$$dP_r = \bar{\rho}[1+\xi(r)]dV_1 dV_2 \tag{1.43}$$

where $\bar{\rho}$ is the mean density. For a random distribution (no clustering), $\xi = 0$ and so, for clustered regions $\xi > 0$ whereas for voids $\xi < 0$.

Wick's theorem For a Gaussian random field, any joint distribution of local densities if Gaussian distributed, meaning that any ensemble average of product of Gaussian random fields ($\langle \delta_1\delta_2\ldots\delta_n\rangle$) can be obtained by the product of ensemble averages of pairs :

$$\langle \delta_1\delta_2\ldots\delta_n\rangle = \sum_{\rm allpairs}\prod_{ij}\langle\delta_i\delta_i\rangle \tag{1.44}$$

In practice, we obtain the ensemble averages of pairs by connecting up all possible pairs of the field and writing down the covariance matrix of each pair using :

$$\langle\delta_i\rangle = \int \delta_i P(\delta|\mu,C)d\delta_i \tag{1.45}$$

where for the initial density field, we have $\mu_i = 0$ for all i. By application of the Wick's theorem, any odd moments (e.g. the third $\langle\delta_1\delta_2\delta_3\rangle$, the fith, ...) can be written as the product of $\langle\delta\delta\rangle$ and $\langle\delta\rangle$. Given that $\langle\delta\rangle$ by statistical homogeneity, all odd moments are found to be zero. So for a GRF with mean 0 and covariance matrix C, all even moments can be written in terms of the two-point, meaning that all the statistical properties are encoded in $\xi(r)$ or $P(k)$.

1.1.3.3 Primordial density power spectrum

One of the key observables that are predicted as a result of inflation is the primordial density power spectrum. Because the inflation assumes a gravitational potential $V(\phi)$ that is invariant under time translation, the universe is expected to be the same on average under the transformation $t \to t + \Delta t$. In terms of what it means for the power spectrum, we need to express the amount of fluctuations that acts at a given scale and states that these fluctuations are constant. It corresponds to primordial perturbations with a **scale-invariant** power spectrum of $P_\phi(k) = A_s k^{-3+(n_s-1)}$ where A_s and n_s are respectively the scalar ampltiude and spectral index of the primordial power spectrum. Indeed, for $n = 1$, the variance of ϕ at a spatial position can be divided into equal amounts of perturbations per constant logarithmic interval in k-space. The relevant quantity to express this scale-invariance is called the **dimensionless power spectrum** [5] $\Delta^2_\phi(k)$ which measures the variance in the density field per logarithmic k interval :

$$\Delta^2\phi(k) = \frac{\delta\sigma^2}{\delta \ln k} \simeq \text{constant} \tag{1.46}$$

The potential ϕ is related to the density perturbation $\delta = \delta\rho/\rho$ by the Poisson equation :

$$\frac{k^2}{a^2}\phi_k = 4\pi G\rho\delta_k \tag{1.47}$$

A similar definition exists for the dimensionless density power spectrum $\Delta^2(k) = \frac{k^3 P(k)}{2\pi^2}$ which measures the contribution of the fluctuations per logarithmic interval at a given k to the variance in the matter density fluctuations. It is therefore clear that for the density perturbation inflation predicts,

$$P_\delta(k) = A_s k^{n_s} \tag{1.48}$$

$$\Delta^2(k) = A_s k^{n_s+3} \tag{1.49}$$

The spectral index n_s is well-constrained by the CMB observations, in particular, the latest results from [19] gave $n_s = 0.9655 \pm 0.0062$. The most relevant constraints put by CMB will be summarized at the end of section 1.1.
A last key prediction of the simplest models of inflation is that the perturbations of the matter and radiation number densities are equal, i.e. they are **adiabatic**.

1.1.4 Structure formation in the cold dark matter paradigm

As time evolves from the Big Bang, the universe passes through the epoch of inflation era, the synthesis of primordial elements and the radiation decoupling until the matter domination that will allow over-dense regions to collapse and form the structures we observe today.
During the radiation-dominated era, the growth of structure is suppressed by the tight interaction of photons and matter. As a result, matter cannot collapse and density enhancements at that period grow slowly until recombination. The amplitude of the k-modes that enter the horizon before the matter-radiation equality ($k < k_{eq}$) is therefore suppressed compared to larger k-modes that enter the horizon after and that can grow faster. This

5. The scale-invariant power spectrum defined by equation 1.46 is also called the Harrison-Zel'dovich spectrum.

suppression of amplitudes for the short k-modes are taken into account by the **transfer function** $T(k)$ which describes the evolution of the density field perturbations through decoupling. It thus relates the primordial power spectrum just after inflation to the linear power spectrum just after recombination :

$$P_L(k,z) = A_s k^{n_s} T^2(k,z) = A_s k^{n_s} \begin{cases} 1 & \text{for } k \ll k_{eq} \\ \left(\frac{k_{eq}}{k}\right)^2 & \text{for } k \gg k_{eq} \end{cases} \tag{1.50}$$

where $P_L(k,z)$ is the **linear matter power spectrum** after recombination. The transfer function depends on cosmological parameters and can be calculated using numerical codes like CMBFAST [35], its successor CAMB [36] or CLASS [37]. In this thesis, we use the CAMB package.

During the matter-dominated era, the expansion of the universe counterbalances the gravitational attraction of matter. Only if there is enough matter for the gravitational interaction to overcome the expansion will allow density enhancements to collapse and grow. At this stage, it is worth remembering that matter is composed of baryons and cold dark matter. Although the microscopic nature of dark matter is unknown, we know dark matter decoupled to the baryon-photon plasma before recombination and start collapsing earlier. [38] were the first to suggest that the global structure formation process must be in two stages : *structure formation* in which dark matter collapse first to form potential wells and then *galaxy formation* where baryons condense within these potential wells defined by the collisionless collapse of dark matter halos. This simplifies the problem in many ways since the complex fluid mechanical and radiative behaviour of the gas associated with the formation of baryonic structures can be ignored initially.

1.1.4.1 Bottom-up formation process

The standard picture for the formation of structures lies on a fundamental mechanism called **gravitational instability** first proposed by G. Lemaitre. In this scenario, large-scale structures are the result of a competition between gravity, which tends to amplify local density fluctuations, and pressure and expansion which suppress them.
The relevant scale in this mechanism is the *Jeans length* which is the scale of fluctuation where pressure and gravitational forces are the same :

$$\lambda_J = \frac{c_s}{\sqrt{G\rho}} \tag{1.51}$$

Depending on the order in which structures undergo collaspe, there are two scenario for structure and galaxy formations :

— **bottom-up** : structures of lower characteristic scales undergo gravitational collapse first before merging into larger structures

— **top-down** : structures of larger characteristic scales undergo gravitational collapse before fragmenting into smaller structures

Dark matter has thus an important impact on structure formation that depends on its nature : cold (non-relativistic) or hot (relativistic). Using numerical simulations of the universe, it has been shown that in the CDM scenario, the smallest structures collapse first and then hierarchically merge into larger structures (i.e. in a bottom-up process). If instead, we consider a scenario for structure formation based on hot dark matter, small-scale structures will end up forming considerably later than in the CDM case. The effect of cold vs hot DM on structure formation is shown in figure 1.10.

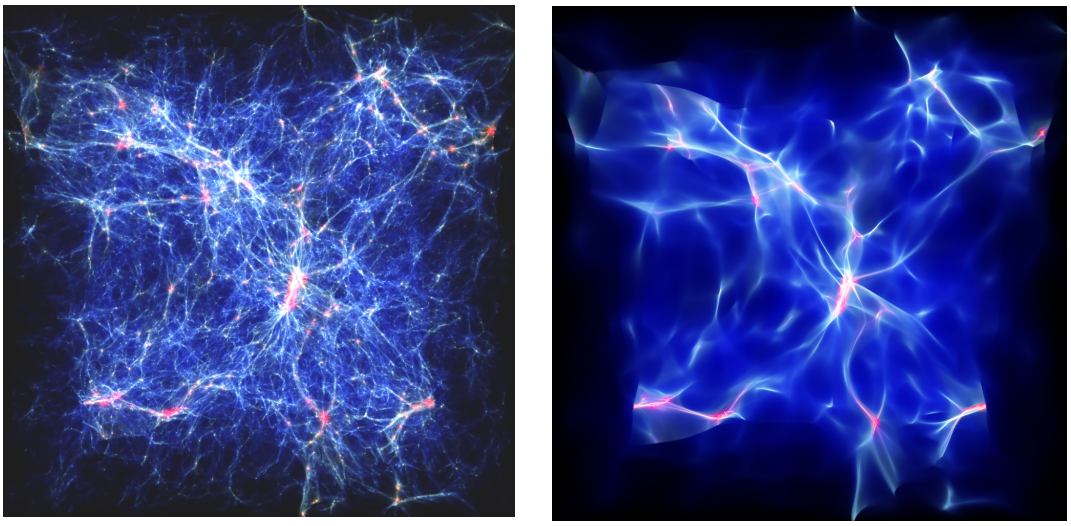


FIGURE 1.10 – Structure formation in the CDM (left) and in the HDM scenario (right). TheHDMr hypothesis is now excluded by the measurements of the matter power spectrum in agreement with a CDM model. From [39].

Measurements of the matter power spectrum using five tracers revealed that its shape was well-fitted by assuming a **cold (non-relativistic) dark matter** [40]. The left panel of figure 1.11 shows the matter power spectrum measured at $z = 0$ while the right panel is an illustration of the shape of the matter power spectrum depending on whether dark matter is assumed to be non-relativistic (cold), relativistic (hot) or a mix of both contributions. The cutoff in the matter power spectrum at small scales (large k) for the HDM scenario is the consequence of the smoothing of structures we mentioned which depends on the dark matter mass. Without a CDM component, the epoch of galaxy formation would occur substantially later in the universe than observed.

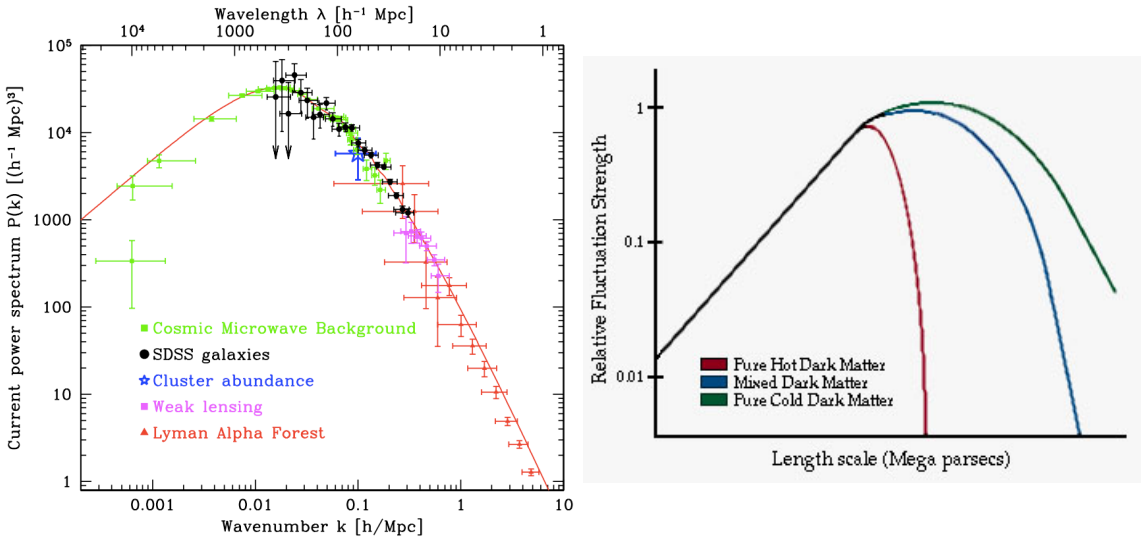


FIGURE 1.11 – left : Matter power spectrum measured at $z = 0$ using five probes. From [40]. Right : Ilustration of the different matter power spectra depending on whether dark matter is assumed ot be cold, hot or a mixuture of both.

1.1.4.2 Cold dark matter

Observational evidence for dark matter dates back the 1930's with the work led by Jan Oort to account for the orbital velocities of stars in the Milky Way [41] and then by Fritz Zwicky to explain the dynamics of galaxies in the Coma galaxy cluster. [6] Zwicky was the first to inferred the total mass of the cluster by measuring the radial velocity dispersion of its galaxies and then applying Newton's law of gravity [42]. He found that the total mass is much higher than the visible mass (gas and stars). In other words, if the total mass of the cluster was in luminous matter only, most of the galaxies would have escaped given their high velocity dispersion. Departures from the predictions of Newtonian gravity also became apparent at galactic scales with the measurement of rotation curves of spiral galaxies [43] which is illustrated in figure 1.12. In order to explain the constant velocities at high radii of the rotation curve, we need to assume the presence of a non-visible massive halo around the galaxy, called a *dark matter halo*. Its mass can be about 10-100 times the visible mass of the galaxy. Further studies have been proposed using the LOS velocities and proper motion of satellite galaxies to determine the galactic halo mass [e.g. 44, 45].

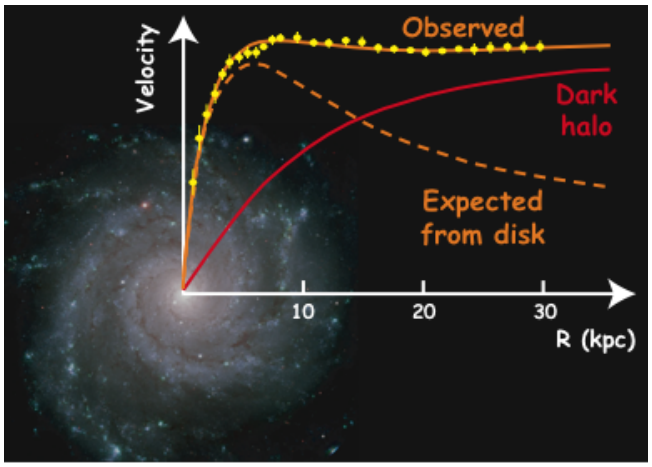


Figure 1.12 – Rotation curves of stars in a spiral galaxy. The prediction (dashed orange curve) from Newton's law of gravity applied to the visible disk does not explain the observed constant behavior at large radii except if we assume an additional mass coming from the dark matter halo around the galaxy.

Since then, additional evidence includes the BBN predictions that $\Omega_b \approx 0.04$ and dynamical measurements that imply $\Omega_m \sim 0.3$. This is confirmed and very well-constrained by CMB measurements with $\Omega_b = 0.04$ and $\Omega_m = 0.3$ [19]. Observational evidence also includes the X-ray emission from the very hot gas at the center of clusters but also the distortions of images due to gravitational lensing in the presence of massive objects. Both effects are illustrated in figure 1.13.

Candidates for dark matter Observations tell us that dark matter makes up 80% of the matter content in the universe, it interacts very weakly and, at least, gravitationally, with ordinary matter. It must be non-relativistic to avoid smoothing out small structures, and it is neutral. Current searches for dark matter focus on non-baryonic can-

6. Clusters of galaxies are the largest gravitationally bound system known in the universe, they contain about 10-1000 galaxies.

FIGURE 1.13 – Composite optical image of the Bullet cluster, with X-ray in pink and weak gravitational lensing in blue (credit : NASA / STScI ; ESO WFI ; Magellan / U. of Arizona). Right : Gravitational lensing manifest near the 0024+1654 cluster (credit : HST), distorting the light rays from a background galaxy, shown as the stretched blue streaks.

didates, which imply new particles beyond the standard model of particle physics. Two leading particle candidates are the neutralino and the axion. The neutralino is an example of WIMP which is predicted in supersymmetry, an extension of the standard model of particle physics. Axions are scalar particles that have been introduced originally [46] to solve another problem in quantum chromodynamics (QCD). Their mass is very low compared to than the one of WIMP (it can range from $10^{-3} - 10^{-4}$ eV for QCD axion to 10^{-22} for ultra-light axion-like DM) and they are supposed to be neutral and stable. Compact objects such as solar-mass black holes are also studied as candidates for dark matter, so that the parameter space for the dark matter mass ranges from 10^{-22} eV to 10^{67} eV. Modifications of Newtonian gravity at the galactic scales as in Modified Newtonian Dynamics theories [MOND, 47] are also considered as a possible explanation for dark matter.

1.1.5 Late-time cosmic acceleration

A last, but not least, unknown in our understanding of the history of the universe is related to the discovery of its late-time acceleration. After the discovey of the expansion of the universe by Hubble, people removed the cosmological constant that was introduced by Einstein to obtain a static evolution for the universe from the Einstein equations. At the end of the 1990s, the HBB model was largely accepted by the community and one important challenge was to measure the expansion of the universe as precisely as possible. People believed that the speed of expansion should decrease with time due to the attractive gravitational interactions between galaxies, so they introduced the deceleration parameter q_0 defined as :

$$q_0 = -\left[\frac{\ddot{a}a}{\dot{a}^2}\right]_{t=0} \tag{1.52}$$

with the minus sign to make q positive for a decelerating universe ($\ddot{a} < 0$). The deceleration parameter was introduced to describe the relation between the distance infered from luminosity, d_L, and the redshift such that :

$$d_L(t_0) = \frac{c}{H_0}\left(z + \frac{1}{2}(1 - q_0)^2 z^2 + \cdots\right) \tag{1.53}$$

Using the Friedmann equation 1.22, the deceleration parameter can be measured through $\Omega_{m,0}/2-\Omega_{\Lambda,0}$. In 1990, the analysis of the large-scale structure correlations of the Automated Plate Measurement (APM) survey using the UK Schmidt Telescope (UKST) favored a scenario in which $\Omega_{m,0} = 0.3$ [48], following by the measurement of the actual baryon fraction in clusters of galaxies compared to the baryon density from BBN in 1998 [49]. However, because inflation predicts a flat universe which is in agreement with CMB observations associated with low redshifts measurements, so $\Omega_k = 0$ and thus $\Omega_{tot,0} = \Omega_{m,0} = 1$, which was supported by other measurements of the mass density [50, 51]. So in 1997, there were two opposite scenarios : a sub-critical universe ($\Omega_{tot} < 1$) with $\Omega_{m,0} = 0.3$ or a critical universe dominated by matter ($\Omega_{tot,0} = \Omega_{m,0} = 1$). But the combination of CMB, LSS, age of the universe and inflationary models led some cosmologists to prefer models with a cosmological constant in order to conciliate a critical universe with $\Omega_{m,0} = 0.3$. [e.g. 52, 53]. Conflict between high values of the Hubble constant (whose inverse gives an estimate of the age of the universe) and the ages of globular clusters also favored a cosmological constant [e.g. 54, 55].

In 1998, two independent teams, the Supernova Cosmology Project [56] led by Saul Perlmutter and the High-z Supernova Search Team [57] led by Brian Schmidt started a program of Type Ia supernovae to measure precisely q_0 up to higher unexplored redshifts ($z \approx 0.7$).
Type Ia supernovae are explosions of white dwarfs that are accreting matter from a companion star until they reach a critical mass at which a runaway nuclear detonation takes place. The star explodes and briefly has a luminosity of about 5×10^9 times that of the Sun. The extreme brilliance of the explosions, combined with the fact that the rate at which they brighten gives a measure of their true brightness, make Type Ia supernovae standardizable candles [7] (objects with a known or derivable luminosity) for measuring the distance to remote galaxies. In addition, the redshifts of the light from the supernovae give access to the recession velocities of the galaxies in which they occur. However, type Ia SNe are not perfect standard candles and the recognition that the peak luminosity of SNe was tighly correlated with the shape of the light curve [e.g. 58] played a critical role, reducing the intrinsic distance error per SNe to ∼10%. The researchers expected to find that the rate of cosmic expansion was greater at greater distances, corresponding to earlier times in the universe. As a consequence, SNe should appear brighter to us as their light has less distance to travel compared to a constant rate of expansion. This is what figure 1.14 illustrates.

This would have been the case if the only component affecting the expansion rate were the mutual attraction of gravity between all the matter that the universe contains. However, their observations showed instead that the SNe appeared fainter, meaning that the rate of cosmic expansion is greater today than what would be expected with deceleration. In other words, 'normal' gravity is not the only influence at work at cosmic scales and another phenomenon, which became known as **dark energy**, is causing the universe to expand at an ever-increasing rate.
So, none of the scenarios mentioned above was in agreement with the observations but [57] and [56] showed that a cosmological model with a **cosmological constant** Λ could explain the SNe data. When combined with CMB observations, [56] even showed that data are consistent with a flat universe where $\Omega_{\Lambda,0} = 0.7$ (see figure 1.15).

7. Cepheid variables, that were studied by Henrietta Leawitt and then used by Hubble to show the expansion of the universe (see section 1.1.1.2), are highly effective standard candles as their period-luminosity relationship is a robust signature.

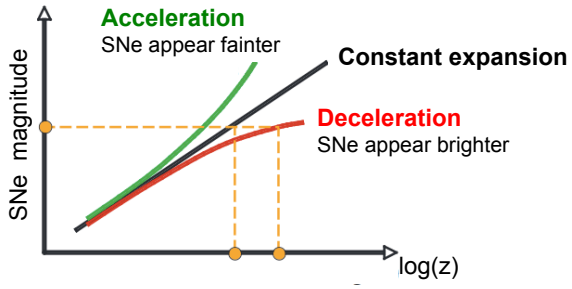


FIGURE 1.14 – Illustration of the relation between the apparent brightness of type Ia SNe and the rate of cosmic expansion. If distances were growing faster in the past, type Ia SNe should have appeared brighter than if the expansion rate was constant.

In the next section, we examine the different models to account for this late-time acceleration of the expansion of the universe.

1.1.5.1 Dark energy models

A cosmological constant It is the mathematically simplest solution to the late cosmic acceleration where we remind the action of GR (with the convention $h = c = 1$) :

$$S_{\rm GR} = \int d^4x\sqrt{-g}\left(\frac{-M_P^2}{2}R - \Lambda\right) + S_{\rm matter}(g_{\mu\nu}, \psi) \tag{1.54}$$

where g is the metric, M_P the Planck scale, R the Ricci scalar and Λ the cosmological constant. We interpret Λ as a new energy component whose properties are constant in space and time. We already said that for an ideal fluid with energy density ρ and pressure p, the effective gravitational source term in GR is $((\rho + 3p)/c^2)$. Dark energy as a cosmological constant is characterized by the equation of state

$$w = \frac{p_{DE}}{\rho_{DE}} = -1 \quad \text{where} \quad \rho_{DE}(z) = \rho_{DE,0}(1+z)^{3(1+w)} = \rho_{DE,0} \tag{1.55}$$

which gives $(\rho + 3p)/c^2 < 0$. So a form of energy that is constant in space and time ($w_\Lambda = -1$) must have a repulsive gravitational effect [e.g. 59]. On the microscopic scale, we could interpret Λ as the gravitational signature of the quantum vacuum energy. But there is a huge problem of magnitude : quantum mechanics is supposed to be valid until the Planck scale M_P which sets for the vacuum energy $\rho_\nu \sim M_P^4/16\pi^2$. Estimations of ρ_ν are about 10^{120} orders of magnitude larger than the observed value for the cosmological constant [60]. Refinements in the calculations by allowing new physics before the Planck scale, like supersymmetry, decrease the difference but it is still about 50 orders of magnitudes [e.g. 60, 61] We thus need to suppose the existence of a mechanism that cancels the vacuum energy in order to keep only what is necessary to account for the late-time cosmic acceleration. As for the flatness problem, it requires to invoke a *fine-tuning* of the cosmological parameters.

Evolving dark energy A natural extension to a cosmological constant is a field with negative pressure to have a repulsive gravitational effect but whose energy density varies with time [e.g. 62, 63]. By analogy with the inflaton field, one can consider a scalar

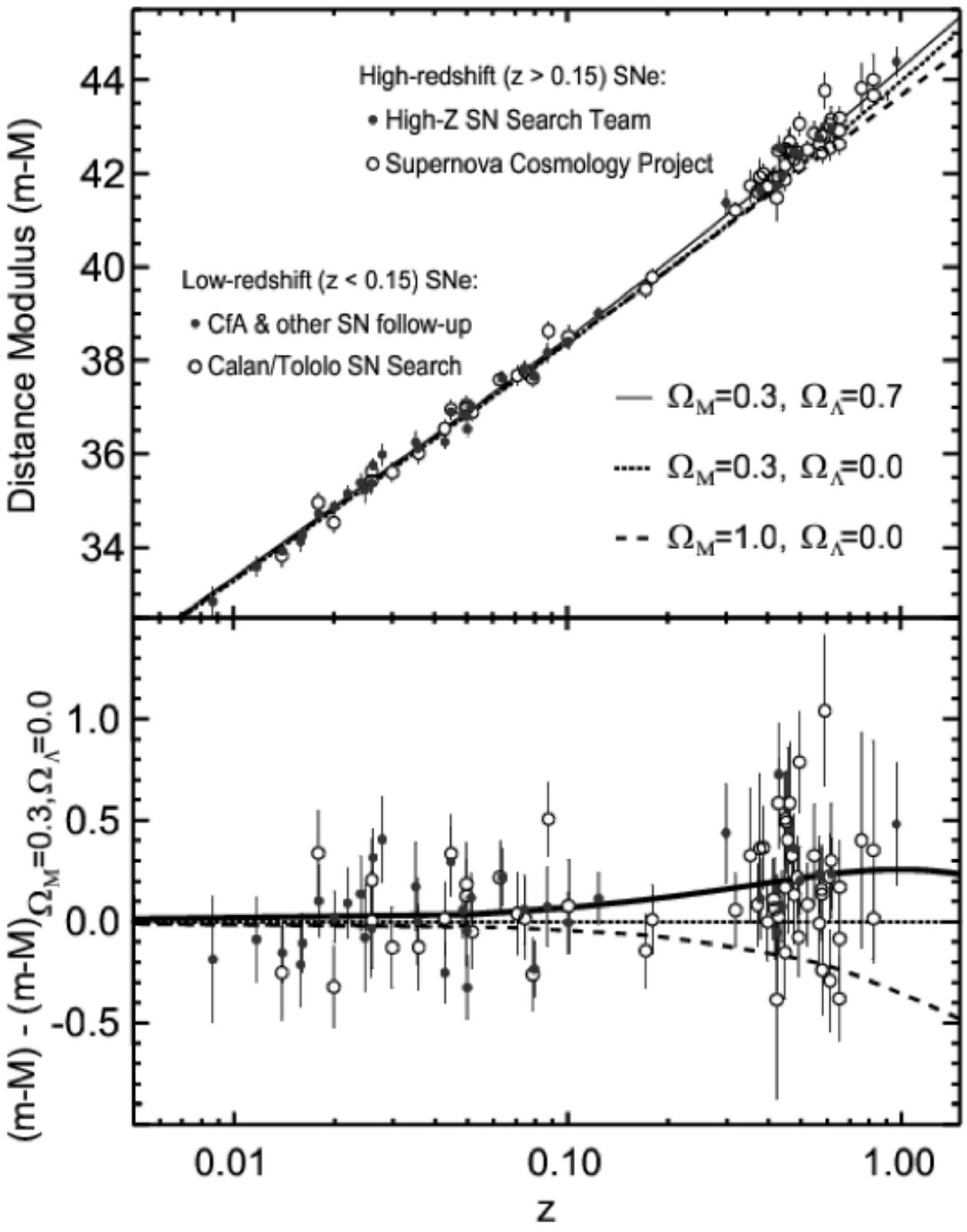


Figure 1.15 – Discovery of the late-time acceleration of the expansion using type Ia SNe by two independent teams, High-Z SN Search Team and Supernova Cosmology Project. Low-redshift measurements are also shown, as well as cosmological models with different energy contents. The one with the cosmological constant density $\Omega_\Lambda = 0.7$ best reproduces the data.

field ϕ potential $V(\phi)$ which enters the action :

$$S = \int d^4x\sqrt{-g}\left(\frac{-M_P^2}{2}R - \frac{1}{2}(\partial\phi)^2 - V(\phi)\right) + S_{\text{matter}}(g_{\mu\nu}, \psi) \tag{1.56}$$

We will refer to this model as **quintessence** which corresponds to a sub-class of scalar field DE for which the energy density is dominated by $V(\phi)$ [64]. Quintessence models assume no coupling between ϕ and matter and behave as a perfect fluid with equation of state given by :

$$w = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)} \tag{1.57}$$

We will see in section 1.2.2.3 that current observations restrict w to be very close to -1. It means that the evolution of the field ϕ is dominated by its potential so that the present situation of cosmic acceleration in quintessence models is similar to a period of inflation. Moreover, the quintessence equation of state is generally time-dependent. A useful two-parameter model to account for the time-evolution of w is [65, 66] :

$$w(a) = w_0 + w_a(1-a) \tag{1.58}$$

which fits many scalar fields expansion histories. This therefore leads to the most commonly used description of dark energy using three parameters : Ω_{DE}, w_0, w_a where the energy density is given by

$$\rho_{DE} = \rho_{DE,0}a^{-3(1+w_0+w_a)}e^{-3w_a(1-a)} \tag{1.59}$$

However, none of these models adresses the *cosmological constant problem* of why the vacuum energy would be so small and one needs to specify the potential $V(\phi)$ and initial conditions for ϕ so that it looks like a fine-tuning is also required to a certain extent.

1.1.5.2 Alternative-gravity models

An alternative to introducing a new energy component is to modify GR itself on cosmological scales. This approach includes a large variety of modified gravity (MG) models. Among the existing explorations to go beyond GR, we can distinguish theories that rely on an additional field (we will focus on *scalar-tensor theories*), *extra-dimensions theories* and *massive gravity theories* where the gravitational interaction is associated with a massive particle, called the graviton. The frontier between DE and MG models is subtle, one key feature of MG models is the introduction of a degree of freedom which is associated with a fifth force so that a screening mechanism is required at small-scales to recover GR.

Scalar theories Scalar theories represent any consistent theory of a metric interacting with a scalar field. In order to avoid unphysical solutions, we usually restrict the studies to theories which have equations of motion of second-order in time derivatives at most. A natural modification of GR is to replace the Ricci scalar R in the action given by equation 1.54 by some function of R. Such a specific case is referred to as $f(R)$ **gravity** and we will describe it in more detail as an illustrative example. First, let define the *Einstein frame* where the Einstein form of the action in GR given in equation 1.54 is preserved and modifications of gravity correspond to additional terms in the action. In contrast, the *Jordan frame* is the physical frame in which we work from the beginning, where energy is conserved and where matter falls along the geodesics. In this frame, the action of any alternative gravity model will be different from the action of GR. The two

frames are related through a conformal coupling (in geometry, a conformal transformation preserves angles locally) :

$$\tilde{g}_{\mu\nu} = e^{-\alpha_i(\phi)/M_P} g_{\mu\nu} \tag{1.60}$$

where $\alpha_i(\phi)$ are the coupling between the massive field ϕ and matter fields. The Einstein frame is used to show the modifications of GR better, but when compared to observations, all equations must be written in the Jordan frame.
The action of $f(R)$ theories in the Jordan frame is :

$$S_{\mathrm{f(R)}} = \int d^4x\sqrt{-g}\frac{-M_P^2}{2}f(R) \tag{1.61}$$

where $f(R)$ is a function to be choosen, for instance [67] showed that a function $f(R) = R - \mu^{2(n+1)}/R^n$ with $n > 0$ and μ a constant. Because the modifications of gravity must be weak in order to agree with the experimental tests of GR, a Taylor expansion could also work : $f(R) = a_0 + a_1R + a_2R^2 + \cdots$ where $a_0 = 0$ to avoid a cosmological constant, $a_1 = M_P^2/2$ to recover GR and a_iR^i would be small compared to a_1R. Using the time arrival of radio waves transmitted from the Cassini space probe, the Cassini mission put the best constraints on a_2 with $|a_2| < 1.2 \times 10^{18}\mathrm{m}^2$ [68].
One can also write the action in the Einstein frame,

$$S_{\mathrm{f(R)}} = \int d^4x\sqrt{-\tilde{g}}[\frac{-M_P^2}{2}\tilde{R} - \frac{1}{2}\tilde{g}^{\mu\nu}(\tilde{\nabla}_\mu\phi)\tilde{\nabla}_\nu\Phi - V(\phi)] + S_i(\chi_i, e^{-\kappa\alpha_i(\phi)\tilde{g}_{\mu\nu}}) \tag{1.62}$$

For $f(R)$ theories, the coupling is universal and linear : $\alpha = \sqrt{2/3}\phi$.
Varying the action defined in equation 5.45 w.r.t. the metric $g_{\mu\nu}$ yields the modified Einstein equation that contains a new degree of freedom $f_R = df/dR$ which is sometines called the scalaron. R can thus be expressed as a function of f_R. For the functional form of $f(R)$ that are viable with experimental tests of GR and current cosmological observations, [69] showed that the quasi-static approximation (QSA) is valid which means that the time derivative of the scalar field can be ignored. In this limit, the scalaron equation is [e.g. 70] :

$$\nabla^2 f_R = \frac{1}{3}a^2[R(f_R) - \bar{R}] + 8\pi G(\rho_m - \bar{\rho_m}) \tag{1.63}$$

where ∇^2 is the 3D gradient operator and the overbar represents the background value of a quantity. The Poisson equation which governs the Newtonian potential Φ is also modified accordingly, and in the QSA it gives

$$\nabla^2\Phi = \frac{16\pi G}{3}a^2(\rho_m - \bar{\rho_m}) + \frac{1}{6}a^2[R(f_R) - \bar{R}] \tag{1.64}$$

So the scalaron can have three effects on cosmology : (i) the background expansion of the universe may be modified because of new terms in the Einstein's equations, (ii) the modification of Poisson equation can change the matter clustering and growth of density perturbations and (iii) the matter perturbation potential Φ and the potential Ψ are different and related by f_R (while in GR, $\Phi = \Psi$). The modifications of the Newtonian gravity can be seen as a fith force which becomes dominant over gravity at large scales. We can distinguish two regimes depending on f_R :

— when $|f_R| \ll 1$, equation 1.63 gives $R \sim -8\pi G\rho_m$ and the second term in equation 1.64 becomes $-8\pi G(\rho_m - \bar{\rho_m})$, so equation 1.64 reduces to the Poisson equation in GR.

— when $|f_R| \gg 1$, we have $|R(f_R) - \bar{R}| \ll 8\pi G|\rho_m - \bar{\rho_m}|$, so equation 1.64 reduces to the Poisson equation in GR with G rescaled by a factor 4/3.

This factor of 4/3 is the maximum enhancement of gravity in $f(R)$ models, it does not depend on the specific functional form of $f(R)$. By chosing a $f(R)$, we set when and on which scale the enhancement factor changes from 1 to 4/3. Because of this 4/3 factor on the strength of Newtonian gravity, the $f(R)$ models would be ruled out by local tests of gravity. But if the function $f(R)$ is chosen appropriately, the effective potential $V_{\text{eff}}(\phi) = V(\phi) - Q\rho_m\Phi/M_P$, due to the coupling between the scalaron ϕ and matter, has a minimum that depends on the mass of the object which is coupled with the scalaron. The more massive the object is, the closer to small values of Φ the minimum of the potential V_{eff} tends to be and so the smaller the coupling between the scalaron and matter is. This *screening* effect that suppresses the enhancement of gravity in the vicinity of high matter density regions is called the **chamelon mechanism** [71, 72]. It allows a class of models [e.g. 73, 74], among which $f(R)$, to satisfy observational constraints in the Solar System and [e.g 75, 76].

Extra-dimensions theories In 2000, Gia Dvali, Gregory Gabadadze and Massimo Porrati proposed an alternative gravity theory where gravity is allowed to propagate in 5D space-time with an additional spatial dimension [77]. Our universe with $D = 3 + 1$ dimensions would be embedded in a $D = 4 + 1$ space, where the extra dimension, called a brane, is assumed to be flat and infinite. In the Jordan frame, the action of DGP becomes :

$$S_{\text{DGP}} = S_{GR} - \frac{M_5^3}{2}\int d^5X\sqrt{-G_{AB}}R_5 \tag{1.65}$$

where M_5 is the Planck mass in a 5D space, $X = (x^\mu, y)$ where y is the position along the extra dimension, G_{AB} is the metric of the 5D space and R_5 the associated Ricci scalar. In this model, the 4D-term dominates at small distances thanks to a screening effect called the **Vainshtein effect** [78, 79] in order to agree with local tests of gravity. The 5D-term is dominant on cosmological scales and is responsible for the acceleration of the expansion and the transition scale is given by $r_c = M_P^2/(2M_5^3)$. Assuming a FLRW metric for the background the DGP action has two solutions depending on the sign of y [79]. However, even if one of these solutions does explain the late-time acceleration, the framework suffers theoretical instabilities and unphysical behaviors [80, 81]. Moreover, [82] showed that observational constraints rule out the simple 5D-model. A way of agreeing with current constraints is to consider a ΛCDM background with a cosmological constant in addition to the extra dimension. It corresponds to the normal branch of DGP and it is referred as the **nDGP model**. However, since the original motivation to introduce modifications of GR that account for the acceleration of the expansion of the universe without the cosmological constant problem, the nDGP model lost its most attractive feature.

Massive gravity By allowing the *graviton* to be massive, it decreases the range of the gravitional interaction, so as its strength between distant objects. If the graviton is massive with mass m, then its potential is modified as follows :

$$\Phi_N = -\frac{G_N M}{r}e^{-mr} \tag{1.66}$$

where r is the distance to the object. However, theoretical instabilities are associated with the introduction of a massive graviton [83, 84], even when non-linear terms are added using the Vainshtein effect [78], instabilities are still present [85]. [86, 87] proposed a viable theory

that recovers GR locally using a Vainshtein effect, however theoretical problems remain when applying the theory to the FLRW metric [88, 89].

Galileon models Galileon theories rely on an additional scalar field π that satisfies the Galilean symmetry : $\pi \rightarrow \pi + c + b_u x^\mu$, where c is a constant and b_μ a vector. It means that the π field, which is called the Galileon, is invariant under translation, and its gradient is also invariant under translation by a constant. To avoid theoretical instabilities as the ones observed in DGP and massive gravity models, the equation of motion is a function of second derivatives of π at most. Imposing such conditions imply only five non-vanishing Lagrangians where the first one is equal to π itself, the second one is related to the kinetic term and the three remaining depend on derivatives of π in a non-linear way. The first Galileon model was proposed by [90]. One can also add a coupling between the Galileon and matter [91], which can be conformal (proportional to π) or disformal (proportional to derivatives of π) but [92] showed that the conformal coupling to matter was disfavored using CMB, SNe and LSS data. The action of the disformally coupled Galileon in the Einstein frame can be written by :

$$S_{\text{Galileon}} = \int d^4x\sqrt{-\tilde{g}}[\frac{M_P^2}{2}R - \frac{1}{2}\sum_{i=1}^{5}\frac{c_i}{M^{3(i-2)}}L_i - \frac{M_P}{M^3}c_G\partial_\mu\partial_\nu T^{\mu\nu} \tag{1.67}$$

where $L_1 = \pi$ behaves like a cosmological constant, so $c_1 = 0$ is imposed as we look for alternatives to the cosmological constant. The presence of couplings with matter can be interpreted again as a fifth force while the presence of non-linear Lagrangians ($L_{3,4,5}$) are necessary to screen this fifth force at small scales, in the vicinity of massive objects using the Vainshtein effect [78]. Whereas the experimental tests of the Equivalence Principle using analyses of laser ranges to the Moon by the Lunar Laser Ranging [LLR 93] put constraints on the conformal coupling $c_0 < 10^{-2}$, there is no current constraint on c_G in the context of a Galileon model. However, different prescriptions for the Galileon model, including coupling to matter or not and considering the five Lagrangians or a smaller set, have been confronted to data [94, 95, 96, 97, 98, 92].

In section 5.5, we will explore the cosmological implications of our measurements within the ΛCDM+GR model but also beyond witht the examples of quintessence, $f(R)$ and Galileon theories.

Summary : Current status of the ΛCDM model

The model towards which all data are converging is called the ΛCDM model. The name refers to the energy content (see figure 1.1) with ∼5% of baryonic matter and radiation with a radiation component which is now negligible, ∼25% of cold dark matter whose nature is still unknown but particle physics and cosmology are looking for neutral, weakly interacting particles whose mass ranges from ultra-light particles to primordial stellar black holes (from 10^{-22} eV to 10^{67} eV).
The ΛCDM predictions depend on six independent parameters :

Amplitude of the primordial power spectrum	$\ln(10^{10})A_s$	3.089 ± 0.036
Index of the primordial power spectrum	n_s	0.9655 ± 0.0062
Age of the universe	t_H	$13.813 \pm 0.038 \times 10^9$years
Baryon density	$\Omega_b h^2$	0.022222 ± 0.00023
Dark matter density	$\Omega_c h^2$	0.1197 ± 0.0022
Optical depth at reionization	τ	0.078 ± 0.019

In the the hot Big Bang scenario, the universe formed 13.8 billion years ago and underwent several eras as its expansion rate is driven by the energy density according to the Friedman equation 1.22. It started with an hypothetical but necessary inflation era that generates initial density fluctuations due to the quantum fluctuations of the inflaton and predicts a scale-invariant power spectrum. After inflation, the growth of perturbations depends on the energy content of the universe and only sub-horizon perturbations will grow (i.e. perturbations whose typical scale is below the Hubble distance $\lambda_H \propto H^{-1}$). In the radiation-dominated era, baryons are coupled to photons which prevents them for collapsing, but dark matter starts forming structures by gravitational instability. Perturbations start growing faster after the matter-radiation equality at $z \sim 3400$ until $z \sim 0.7$ when dark energy becomes dominant, making distances growing faster and prevent structures from undergoing gravitational collapse.
In what follows, we summarize the properties of each component of the universe :

component	equation of state	energy density	scale factor
radiation	$w_r = 1/3$	$\rho_r \propto a^{-4}$	$a \propto t^{1/2}$
matter	$w_m = 0$	$\rho_m \propto a^{-3}$	$a \propto t^{2/3}$
cosmological constant	$w_\Lambda = -1$	$\rho_\Lambda \propto a^0$	$a \propto e^{H_0 \Omega_\Lambda^{1/2} t}$

1.2 Probing cosmic acceleration with large-scale surveys

Understanding the large-scale structures of the universe is one of the main goals of modern cosmology. Studies of the clustering of galaxies started with the early ideas of [99] that galaxies could be treated as **biased tracers** of matter clustering. So the existence of structures has been suggested by early observational projects which aimed at mapping the distribution of galaxies, which resulted in a number of discoveries of individual structures such as filamentary bridges between superclusters and large empty regions (voids) on scales of tens of Mpc [100, 101, 102, 103]. Then, technological progress including machine plate scanning and CCDs enabled the construction of automatic wide-field surveys in the 1980s.

1.2.1 Brief history of LSS surveys

In 1982, the state-of-the-art survey was the one carried out by the Center for Astrophysics (CfA) redshift survey [104] which measured the radial velocities of the brightest galaxies in the nearby universe. The map obtained from the first CfA survey [105] is shown in the left panel of figure 1.16, it revealed a variety of structures with galaxies actually appearing to be distributed on surfaces, almost bubble-like, surrounded by voids. Then, between 1985 and 1995, a second CfA survey (CfA2) recorded 18,000 bright galaxies in the northern sky whose map is shown in the right panel of figure 1.16.

In 1989, the results of the CfA surveys reached a milestone with the discovery of the *Great Wall* [106] which is the structure running all the way across between 8 hours and 17 hours RA and 5,000 and 10,000 km.s^{-1} in the right panel of figure 1.16. It was the largest single structure yet detected in a redshift survey and definitely highlighted the existence of a *cosmic web* due to the clustered nature of the distribution of galaxies, compared to a random process. It also provided evidence that the CDM paradigm alone was not able to explain the galaxy clustering of CfA galaxies [107] as well as those identified from the first infrared telescope in space, IRAS [108, 109]. Early measurements of the distortions of the redshift space power spectrum date back $\sim$ 1995 using the IRAS survey [110, 111, 112, 113]. These distortions are caused by the line-of-sight component of

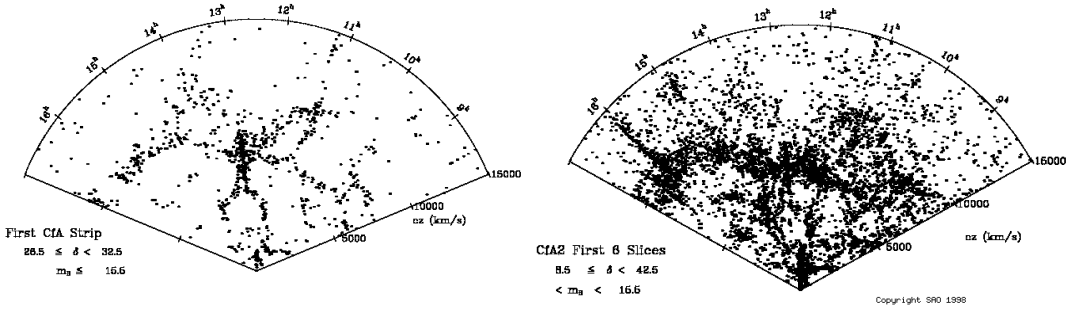


FIGURE 1.16 – Left panel : CfA 1, Right panel : CfA 2 with 18,000 bright galaxies revealing the presence of a *Great Wall*, the structure which is running all the way across between 8 hours and 17 hours in right ascension (RA) and 10,000 km.s^{-1}. Credit : CfA redshift surveys.

peculiar velocities which are included when inferring distances from redshifts. The effect is known as *redshift-space-distorsions* and will be described further in section 1.2.2.2. [114] used the quadrupole-to-monopole ratio to measure the distortions of the power spectrum, then the IRAS Point Source Catalogue Redshift Survey [PSCz, 115] that contains 15,411 redshifts was built and used for RSD measurements [e.g. 116, 117].

In what follows, we propose a brief overview of past, present and future surveys based on the report of the Dark Energy Task Force [DETF, 118] which categorized experiments into four stages and on the review on cosmic acceleration probes by [119]. Figure 1.17 summarizes the most important projects for LSS studies and the most importants weak lensing (WL) programs. Among the LSS experiments, those in red correspond to surveys using the Sloan Foundation telescope (SDSS), those in blue use the facilities at the Anglo-Australian Observatories (AAO) and those in green are ESO programs using the Very Large Telescope (VLT), in Chile. Red contours correspond to spectroscopic surveys and blue contours to photometric ones.

— **Stage I** Stage I experiments include all experiments whose final results were published before 2005.

— **Stage II** Stage II experiments have completed their observations and analyses at the time of writing this thesis. The main projects for LSS include BAO measurements from the Sloan Digital Sky Survey [SDSS-I and II 120, 121, 122] and the 2 degree Field Galaxy Redshift Survey (2dFGRS) of the Anglo-Australian Observatory [AAO, 123, 124]. [125] also led a spherical harmonics decomposition of the 2dF data to measure the distortions of the redshift power spectrum. Other experiments include the SN and WL programs of the CFHT Legacy Survey [CFHTLS, e.g. 126], the ESSENCE SN survey [127] and the SDSS-II SN survey [128]. Complementary analyses include multi-wavelengh studies of local and high-redshift SN (Carnegie Supernova Project [e.g. 129], systematic searches for $z > 1$ SNe with the Hubble Space Telesopce [HST, e.g. 130] and CMB data from the *WMAP* satellite [27]. The main result of Stage II is the consistency of independent data sets and analyses with a spatially flat universe with $\Omega_\Lambda \sim 0.75$ and with an uncertainty in the equation-of-state parameter w of roughly ± 0.1 at the $1 - 2\sigma$ level.

— **Stage III** Some Stage III experiments have completed their observations but not

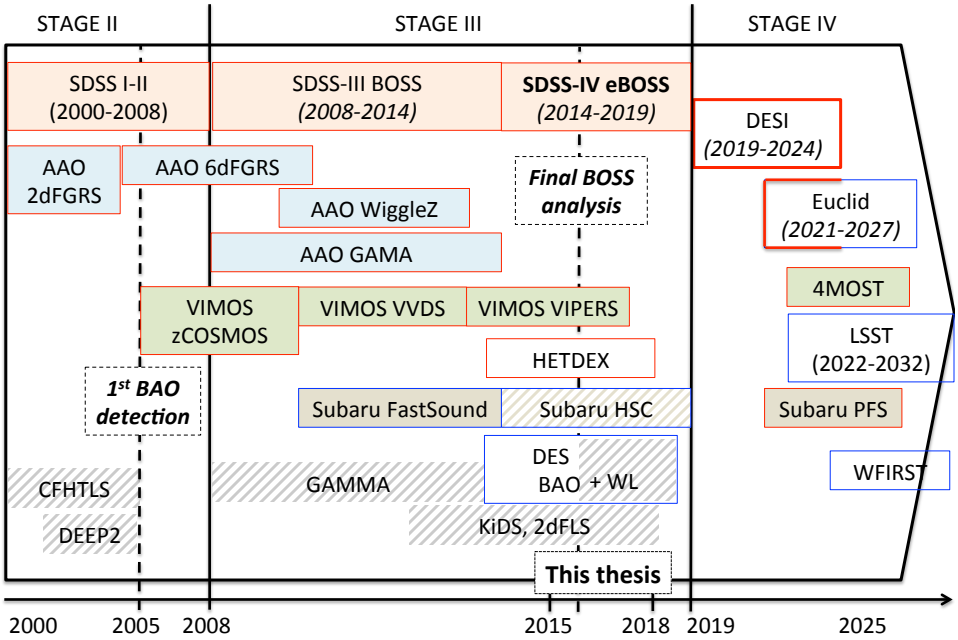


FIGURE 1.17 – Past, present and future main LSS surveys that were categorized by the DETF into four stages. This thesis is using spectroscopic data from the SDSS-IV eBOSS which is one of the Stage III experiments.

necessarily their final analyses, and other projects are still ongoing. Stage III LSS projects include the Baryon Acoustic Oscillation Spectroscopic Survey of SDSS-III [e.g. 131, 132, 133] and the extended BOSS of SDSS-IV [134, 135], the Dark Energy Survey [DES, 136, 137] which uses imaging data only but has also an important WL program, the AAO spectroscopic surveys including GAMA [138], 6dFGRS [139, 140] and WiggleZ [141, 142], the Subaru imaging surveys including FastSound [143] and Hyper Suprime-Cam [HSC, 144] and the ESO VLT programs which focus on targeting galaxies in smaller area on the sky but at higher redshifts with zCOSMOS [145], VVDS [146, 147], VIPERS [148, 149]. The Hobby-Eberly Telescope Dark Energy Experiment [HETDEX, 150] which is using Ly-α emitters at $z \sim 3$ as tracers is also part of Stage III, so as other projects like imaging sly surveys Pan-STARRS [151] and Gaia [152]. Regarding CMB observations, the CMB mission of Stage III is the European satellite *Planck* whose latest results on the most favored cosmological model have been published in 2015 [19].

— **Stage IV** Stage IV experiments are mainly designed to combine many probes. They include deep spectroscopic surveys like the Dark Energy Survey Instrument [DESI, 153, 154], the Subaru Prime Focus Spectrograph [PFS, 155, e.g.] and wide imaging survey like the ESA space mission *Euclid* [156], on the ground the Large Synoptic Survey Telescope [LSST, 157] and the VLT 4MOST [158]. Longer-term projects include the space mission Wide Feld Infrared Survey Telescope [*WFIRST*, 159], the Square-Kilometer Array [SKA, 160, 161] that could enable a BAO survey of $\sim 10^9$ HI-selected galaxies and weak lensing measurements of $\sim 10^{10}$ star-forming galaxies using radio continuum shapes.

This thesis work uses data from SDSS-IV eBOSS. The survey will be described in more

details in chapter 2.

1.2.2 State-of-the-art galaxy clustering analysis

The report of the DETF [118] and [119] consider the BAO method only, here we will also consider the redshift space distortions (RSD) technique combined to the Alcock-Pacskynski effect [AP, 162] that will be explained later in the chapter. Both BAO and RSD are distinct features of the matter clustering one can recover from the distribution of galaxies using photometric or spectroscopic surveys. Because of the lack of precise information in the redshift direction for photometric surveys, most analyses are based on the angular clustering, that is to say integrated along the LOS. In this thesis, we use spectroscopic data from SDSS-IV eBOSS, so clustering analysis in our case refers to the analysis of the power spectrum (equation 1.41) or the two-point correlation function (equation 6.2) of the galaxy density field. When the entire spectrum of frequencies and positions is considered, both estimators contain identical information as one represents the Fourier transform of the other. However, since our Fourier spectral range is finite, they are not exactly the same although we expect a high correlation between results from each method. Moreover, we expect that any potential uncorrected observational or modelling systematics will affect the correlation function and power spectrum differently, so that performing both should yied a more robust final result.

It is also worth remembering that galaxies are **biased tracers** of the matter density, as originally highlighted by [99], which means that the strength of galaxy clustering is different from the predicted one of dark matter. The current understanding is that virialized dark matter halos cluster more strongly than the dark matter distribution as a whole, so the enhanced clustering strength of galaxies could reflect the halos that they inhabit. Depending on the type of galaxies we observe, the mass range of halos that they inhabit will be different. Such a phenomenon can be characterized by the **bias parameter** b defined by :

$$\left(\frac{\delta\rho}{\rho}\right)_{\text{galaxy}} = b\left(\frac{\delta\rho}{\rho}\right)_{\text{darkmatter}} \tag{1.68}$$

where in the most general case, the bias can be a complex function of redshift, scale and local-environment properties. However, on large scales ($\geq\sim 80h^{-1}Mpc$), there are good reasons to beliveve that the bias behaves as a constant multiplicative factor [e.g. 1, 163]. It implies that $\xi_g = b^2\xi_{DM}$. Section 3.3 will describe further our understanding and modelling of the bias, although this relationship is rather considered as an empirical approach to encode complex radiative and cooling processes associated with the physics of baryons and gas.

1.2.2.1 BAO as a standard ruler

The BAO feature in the matter clustering correspond to the imprint left by sound waves from the baryon-photon plasma in the early universe (see section 1.1.2.2). In this plasma, the high photon-baryon pressure resulted in sound waves propagating at the sound speed $c_s \simeq c/\sqrt{3}$ until the baryons are released from the photons at the *drag epoch* at redshift z_d (shortly after the decoupling of photons). By then, the BAO have travelled a distance called the sound horizon at the baryon-drag epoch, $z_d = 1060$, and whose value is :

$$r_s(z_d) = \int_{z_d}^{\infty} \frac{c_s(z)}{H(z)} dz = (99.17 \pm 0.33)\, h^{-1}\text{Mpc} \tag{1.69}$$

where $r_s(z_d)$ is known to 0.3% from Planck data [19]. We use the following notation $r_s(z_d) = r_s$.

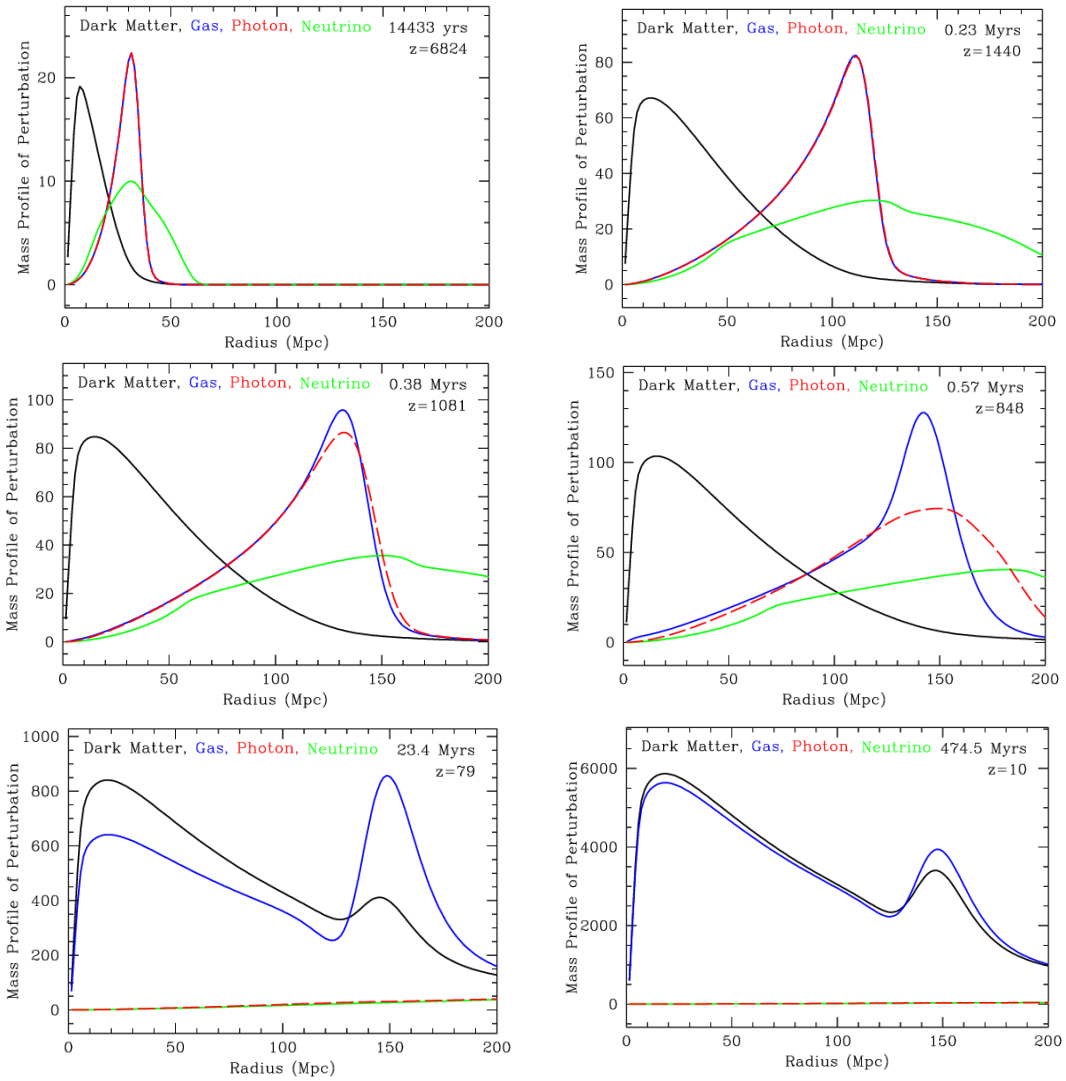


FIGURE 1.18 – Evolution of the radial mass profile for different components of the universe as a function of comoving radius of an initially point-like overdensity located at the origin. Credit : [164]

Figure 1.18 illustrates the radial evolution of the dark matter mass profile (black), gas (blue), photon (red) and neutrino (green) perturbations as redshift decreases, whose steps can be summarized as follows :

1. Snapshot at $z = 6824$: photons and baryons were in thermal equilibrium and the neutrinos already free stream out. Dark matter is not coupled to the rest of the particles and it accumulates in the overall density perturbation around its origin. The snapshot at $z = 1440$ shows that the width of the dark matter perturbation is wider as it is collapsing and accreting additional material in its vicinity.
2. During the recombination (snapshot at $z = 1081$) baryons start decoupling from photons and the mean free path of photons becomes larger than the Hubble horizon so light could cross the universe without scattering. This is why in the snapshot at

$z = 848$, the photon perturbation is moving away from the baryon perturbation and photons free-streamed.

3. In the snapshot at $z = 78$, there are remaining dark matter perturbation around the origin and gas perturbation in a shell of about ∼150 Mpc. But then, baryons start falling into the potential wells formed by dark matter and both interact gravitationally, so that eventually the spherical shell of baryons imprinted itself in the dark matter, as shown in the snapshot at $z = 10$.

So the imprint on the matter clustering corresponds to a characteristic scale with wiggles in the power spectrum at $k \sim 0.07, 0.13, 0.19\,h.\mathrm{Mpc}^{-1}$ and a local enhancement in the two-point correlation function at s $\sim 100\,h^{-1}$.Mpc in comoving units. The identification of the acoustic scale as a standard ruler, first in the CMB and then in the matter clustering dates back to the end of 1990 [165, 166], then a series of papers gave hints of the acoustic feature in the power spectrum using the 2dFGRS and in the two-point correlation function using the dedicated SDSS Luminous Red Galaxy (LRG) sample [167]. However, the first convincing detections of BAO with more thant 3σ significance came in 2005 from the SDSS Data Release 3 which contained ∼ 47,000 LRG in the redshift range $0.16 < z < 0.47$ [SDSS DR3, 120] and from the final 2dFGRS sample that contained ∼220,000 galaxy redshifts [124], as shown in figure 1.19.

Although non-linearities are not dominant at the very large scales of the BAO peak, they move galaxies around by 3-10 Mpc which broadens the peak and degrades our precision on the measurement of its position [168]. [169] developed a model to account for non-linearities by modeling the differential motion of pairs initially separated by 150 Mpc. The large scale correlation function is thus the convolution of the linear correlation function with the distribution of differential motions. So, the acoustic peak is degraded due to bulk flows created by gravitational forces which are sourced by the presence of mass. By mapping the galaxy density field, one can run this backwards and remove most of the non-linearities that degraded the statistical precision. Such a technique is called **reconstruction** and was proposed by [164]. Figure 1.20 shows the monopole of the matter correlation function in redshift space after reconstruction. The black solid line shows the correlation function at $z = 49$ when non-linearities are completely negligible, the blue short-dashed line shows the correlation function at $z = 0.3$. We can clearly see that the acoustic peak has been smeared out. The black dotted line shows the real-space correlation function, and the red and magenta lines show the effects of reconstruction depending on two prescriptions. These reconstructions significantly sharpen the acoustic peak. Reconstruction has been first applied to the SDSS-II DR7 LRG sample at $z_{\mathrm{eff}} = 0.35$ [122] and has been to be particularly efficient in this case, providing a 1.9% distance measurement at $z = 0.35$, decreasing the error by a factor of 1.7 compared to the pre-reconstruction measurement. Now, the technique is always used to infer the best constraints on the BAO distances, when the density of the sample is high enough to reconstruct the density field (otherwise, the technique just adds additional noise to the already noisy data).

However, we do not measure clustering direcly in comoving coordinates, but instead we measure galaxy redshifts and angles from spectroscopic surveys and we infer distances from these. If we use the wrong cosmological model to do this conversion (by wrong, we mean different from the true one of the data that we do not know), then the distances will be wrong and the inferred clustering will contain detectable distortions, independently of RSD (see section 1.2.2.2). This effect is known as the Alcock-Paczynski effect [162] and in association with a standard ruler, it provides a geometrical test to disentangle between RSD and distortions due to the wrong fiducial model.

Figure 1.21 illustrates how we can actually measure distances with the 3D distribution

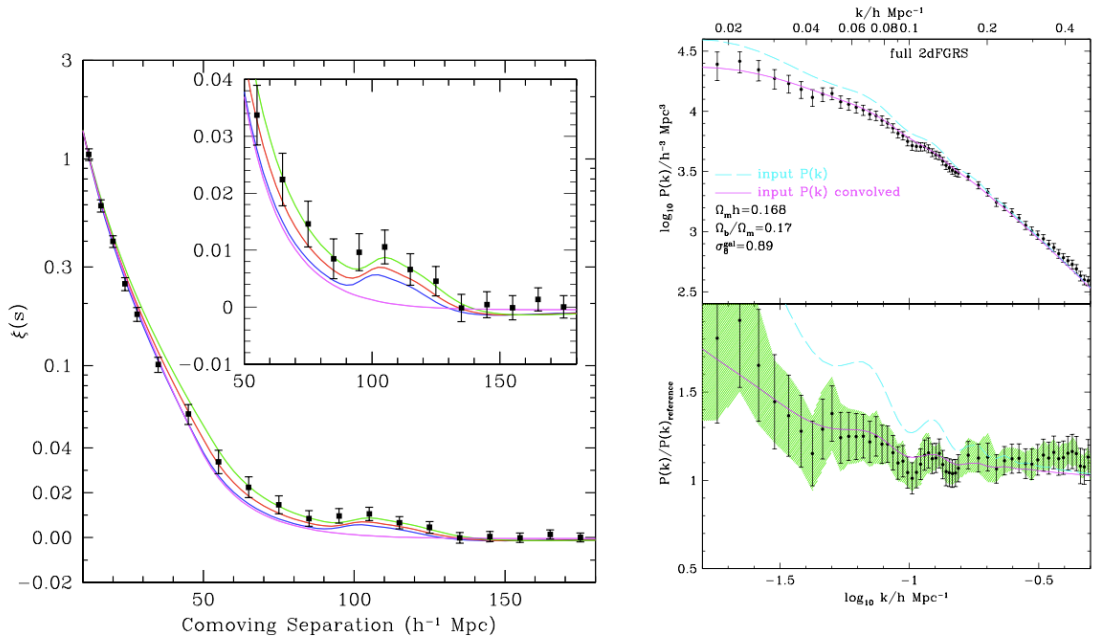


FIGURE 1.19 – First detection of the BAO signal in the two-point correlation function of SDSS LRG (left panel, [120]) and the power spectrum of the 2dFGRS sample (right panel, [124].

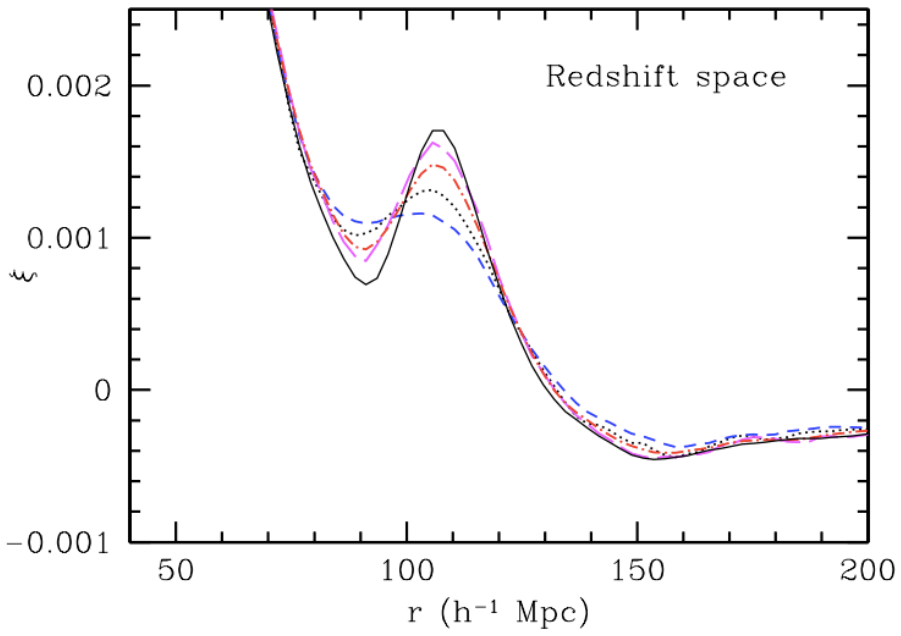


FIGURE 1.20 – Monopole of the matter correlation function in redshift space after reconstruction. The black solid line shows the correlation function at $z = 49$ when non-linearities are completely negligible, the blue short-dashed line shows the correlation function at $z = 0.3$. We can clearly see that the acoustic peak has been smeared out. The black dotted line shows the real-space correlation function, and the red and magenta lines show the effects of reconstruction depending on two prescriptions. From [164]

of galaxies using the AP test and a standard ruler such as the BAO scale. Galaxies which are physically separated by the BAO scale Δl are observed with an angular separation $\Delta s_\perp$ and a radial separation $\Delta s_\parallel$ such that $\Delta s_\parallel = \Delta s_\perp = \Delta l$, where

$$\Delta s_\parallel = \frac{c}{(1+z)H(z)}\Delta z \tag{1.70}$$

$$\Delta s_\perp = D_A(z)\Delta\theta \tag{1.71}$$

Using an anisotropic clustering measurement, the combination of a standard ruler (whose physical size is known or can be predicted) with the Alcock-Paczynski test [162] allows us to measure separately $H(z)$ and $D_A(z)$ [170]. If we consider averaging clustering in 3D over all directions instead, the isotropic shift is sensitive to volume/angle averaged BAO distance :

$$D_\mathrm{V}(z) = \left[(1+z)^2 cz\frac{D_\mathrm{A}^2(z)}{H(z)}\right]^{\frac{1}{3}} \tag{1.72}$$

Although this projection applies to all scales of the clustering signal, the position of the BAO provides the most robust constraint as the BAO is a distinct feature (i.e. a peak) in the correlation function on sufficiently large scales that it is difficult to alter it with non-linear physics. We often introduce LOS and transverse dilation scales which measure departures between the expected BAO position according to the assumed cosmological model and the observed one :

$$\alpha_\parallel = \frac{H^\mathrm{fid}(z) r_d^\mathrm{fid}}{H(z) r_s} \tag{1.73}$$

$$\alpha_\perp = \frac{D_A(z) r_d^\mathrm{fid}}{D_A^\mathrm{fid}(z) r_s} \tag{1.74}$$

where r_s is the comoving BAO scale at the drag epoch measured from CMB anisotropies and the superscript 'fid' denotes the fiducial value of a quantity. And the isotropic shift of the BAO position is related to D_V by

$$\alpha_\mathrm{isotropic} = \frac{D_V(z) r_d^\mathrm{fid}}{D_V^\mathrm{fid}(z) r_s} \tag{1.75}$$

[171] showed that a joint fit of $H(z)$ and $D_A(z)$ led to a degeneracy breaking and to a factor of five improvement on measuring the DE equation-of-state, using the same data set.

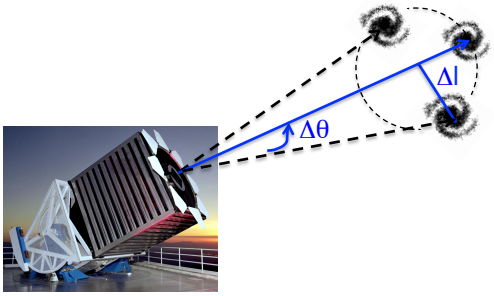


Figure 1.21 –

The challenge of the BAO method is primarily statistical as it corresponds to a weak signal at large scales that requires big volumes of the universe to measure it precisely. At

low redshift ($z \leq 0.5$), the BAO method strongly complements Type Ia SN measurements as it provides an absolute distance scale which is related to the CMB acoustic peaks. The BAO method can also be used at higher redshift where standard astrophysical objects cannot be used as standard candles because we can no longer neglect or model precisely their intrinsic evolution.

1.2.2.2 Linear growth of structure and redshift space distortions

Because of gravitational attraction, the galaxies tend to fall towards high-density regions and flow away from low-density regions, such that the clustering is enhanced in the LOS direction compared to the perpendicular direction. The effect is sometimes referred as the Kaiser or squashing effect [172]. At small scales ($< 10 - 20\,h^{-1}$.Mpc), another effect has to be taken into account : the halo velocity dispersion elongates clusters along the LOS, leading to the "Finger of God" (FoG) effect [173]. Figure 1.22 illustrates how a spherical overdensity appears distorted by peculiar velocities along the LOS when observed in redshift space. Galaxies are undergoing a coherent infall towards a spherical overdensity and the arrows represent their peculiar velocties. At turnaround, the peculiar infall velocity that is just cancelling the global Hubble expansion, appears collapsed to a single velocity in redshift space. Peculiar velocities make large scale galaxy clustering appear anisotropic in redshift space. These anisotropies or distortions are caused by the line-of-sight component of galaxy peculiar velocities affecting the observed galaxy redshifts from which distances are measured. The redshift-space position of a galaxy differs from its real-space position due to its peculiar velocity along the LOS (taken to be the z-axis) :

$$\mathbf{s} = \mathbf{r} + \frac{v_z(\mathbf{r})}{aH}\hat{\mathbf{z}} \tag{1.76}$$

where $\hat{\mathbf{z}}$ is the LOS unitary vector and v_z is the LOS component of the peculiar velocity. On linear scales (above $60 - 80\,h^{-1}$.Mpc), the theory behind the observed redshift space clustering is well developed [172, 174] where the divergence of the peculiar velocity field is directly related to the growth rate of structure, assuming a linear coupling between the density and velocity fields :

$$\theta = \nabla \cdot \mathbf{v} = f\delta_m \tag{1.77}$$

Moreover, [172] showed that the Legendre monopole and the quadrupole of the power spectrum can be related to the bias and the growth rate, and [174] provided the counterpart expressions in configuration space where at first order $\xi_0(s) \propto b^2$ and $\xi_2(s) \propto bf$. The expressions will be provided and explained in section 3.4.1. Thus, if one can determine simultaneously the monopole and the quadrupole, one will be able to extract both $f\sigma_8$ and $b\sigma_8$, where σ_8 is the normalization of the linear power spectrum given by

$$\sigma^2_{8,g} = \int \frac{dk}{2\pi^2} W_8^2(k) k^2 P_g(k,\mu) \tag{1.78}$$

where $W_8(k)$ is the Fourier transform of a top-hat window function of width $8\,h^{-1}$.Mpc. Historically $R = 8\,h^{-1}$.Mpc has been chosen because at $z = 0$, $\sigma_8^2 \approx 1$. In theory, we could derive σ_8 from the amplitude of the matter power spectrum but in practice, we measure the power spectrum of biased tracers, P_g. A change in $\sigma_{8,g}$ would imply a change in the cosmological parameters, such as the bias (amplitude of the clustering of tracers) or the growth rate (amplitude of the anisotropies of the clustering). This degeneracy between parameters cannot be broken using the two-point statistics alone. Therefore, instead of constraining f and b independently, we measure $b\sigma_8$ and $f\sigma_8$.

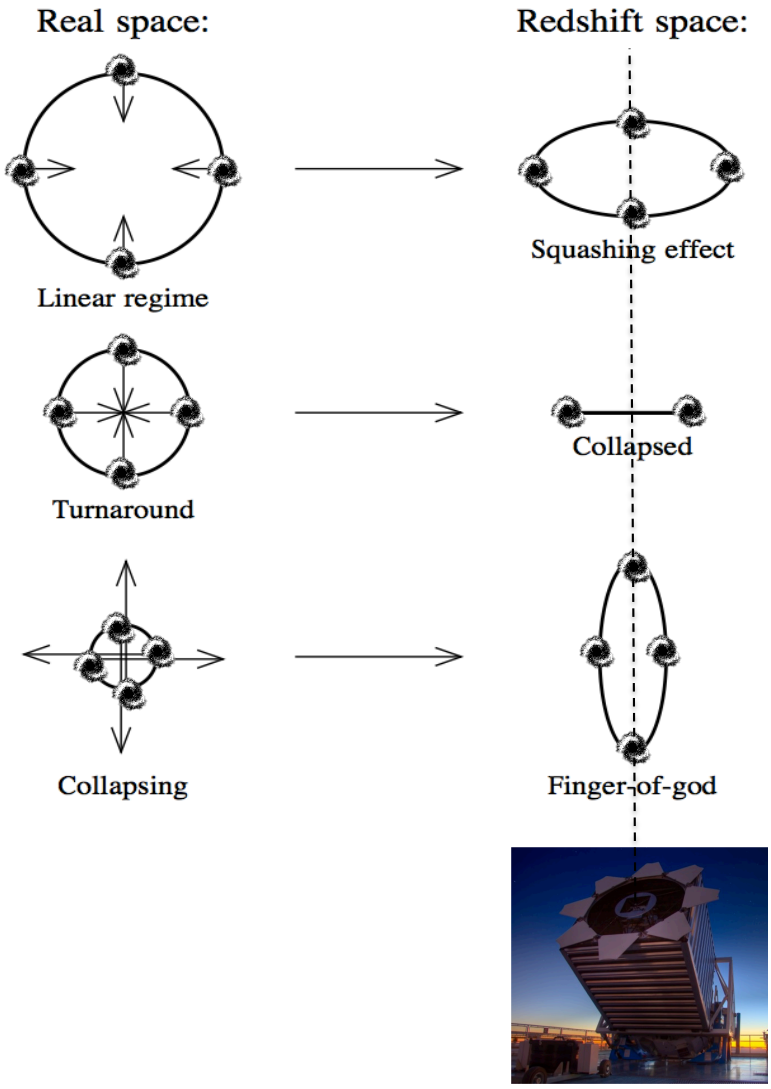


FIGURE 1.22 – Illustration of an overdensity seen in real and redshift space. In redshift space, the projection of peculiar velocities along the LOS is the cause of the anisotropies in the observed clustering.

The strengh of the anisotropy defined by the quadrupole-to-monopole ratio is governed by the distortion parameter $\beta = f/b$ [175] where $f \sim \Omega_{m,0}^{0.55}$ with $\Omega_{m,0}$ is the present matter density. The distortion parameter has been measured very early for a variety of galaxy redshift samples [e.g. 114, 176]. By modeling the full-shape redshift-space galaxy power spectrum, one can extract the parameter combination $f(z)\sigma_8(z)$, the product of the matter clustering amplitude defined in equation 1.78 and the linear growth rate defined in equation 1.77. For a clear review of the physics of RSD, see [177]. In addition to measure β at $z = 0.8$, [178] emphasized that large-scale peculiar velocities were relevant to constrain models of cosmic acceleration. Then, using the data of the VIMOS VLT Deep Survey (VVDS) [146] measured a growth rate measurement consistent with ΛCDM+GR prediction. Since then, RSD have been of great interest as they are a powerful probe to constrain the growth of structure and complementary to BAO distance measurements.

BAO-only vs full-shape clustering analysis Contrary to the BAO technique, RSD studies require to model the full-shape of the two-point correlation function (or power spectrum). Measuring the relative clustering in both LOS and perpendicular directions leads to a measurement of the growth rate of structure, but which is degenerate with the AP effect. By measuring simultaneously $f\sigma_8$ and the anisotropic positions of the BAO (to derive constraints on $H(z)$ and $D_A(z)$), one can disentangle both effects [e.g. 179] and provide a measurement of f which does not depend on the fiducial cosmology assumed to convert redshift and angles into distances.

The key challenge in modeling RSD is to account for non-linear effects that arise from the non-linear evolution of the density and velocity fields, but also from the non-linear and/or scale-dependent bias between galaxies and matter and the non-linear mapping between real to redshift space. The linear theory formalism is not enough even on scales below $50 - 60\,h^{-1}$.Mpc because of a variety of non-linear effects, including the FOG distortions that occur in collasping and virialized regions at small scales. Details on the theoretical framework behind RSD will be presented in Chapter 3.

1.2.2.3 Current constraints

The final constrains from BOSS [132] as part of SDSS-III [131] comes from the analysis of the final data release DR12 with a sample of 1.5 million targets distributed across a footprint of nearly 10,000 deg^2. The analysis was performed using two samples : a sample of LRG up to $z \sim 0.4$ (LOWZ) and a sample of massive galaxies from $0.4 < z < 0.7$, and the redshift range was divided into three bins : $0.2 < z < 0.5$ with $z_{\rm eff} = 0.38$, an overlapping bin $0.4 < z < 0.6$ with $z_{\rm eff} = 0.51$ and $0.5 < z < 0.75$ with $z_{\rm eff} = 0.61$. BAO and full-shape analyses have been performed to measure D_A, H and $f\sigma_8$, and all analyses have been combined to provide a consensus [133]. BOSS also targeted quasars at $z > 2.2$ with the main goal of using them as indirect tracers to study cosmology with the Lyman-α forest. Absorption features from the continuously redshifting Ly-α line in the spectra of distant quasars due to the neutral hydrogen in the high redshift universe can be used to infer the clustering of Ly-α forest flux along single LOS. Note that the BAO measurements using Ly-α forests that are reported in [133] correspond to the analysis of DR11 [180, 181], the results using the final sample DR12 are presented in [182, 183].

Table 1.1 and table 1.2 summarize the constraints from BAO-only and full-shape clustering analyses respectively using various surveys. When no values for D_A and H are reported, it means that $f\sigma_8$ has not been measured with the AP parameters, except for [185] where they performed the full fit but did not report the values of the AP parameters.

Table 1.1 – Precision of the BAO distances from BAO-only analyses using various large-scale structure surveys.

Survey	Reference	z	D_V	D_A	H
6dFGS	[139]	0.106	4.5%	–	–
WiggleZ	[142]	0.44	4.8%	–	–
		0.60	4.5%	–	–
		0.73	3.4%	–	–
SDSS MGS	[184]	0.15	3.8%	–	–
SDSS BOSS galaxies	[133]	0.38	–	1.6%	2.5%
		0.51	–	1.5%	2.6 %
		0.61	–	1.6%	3.0 %
SDSS BOSS Ly-α	[182]	2.4	–	5.8%	3.4%
	[183]	2.33	–	3.9%	2.8%
	combination	2.4	–	3.0%	2.0%

[186] combined the BOSS CMASS sample with the Wiggle Z data in the regions where both survey overlapped and found the following constraints on D_V in the redshift range $0.4 < z < 0.6$: 2.3% for the auto-correlation of CMASS, 3.0% for the cross-correlation and 9.5% for the auto-correlation of WiggleZ galaxies. A direct comparison between the latter and constraints from [142] is not possible as the overlapping redshift range and regions for BOSS+WiggleZ is not the same as the ones for WiggleZ alone.

1.3 Outline of this thesis

In the light of the overview of LSS surveys presented below, current constraints have probed both the redshift range $z < 1$ using galaxies as direct tracers of the matter density field and the high-redshift range $z > 2.1$ using the Lyman-α forests in quasar spectra as indirect tracers of the neutral hydrogen in the intergalactic medium (IGM). In constrast, eBOSS opens up the $z < 2.2$ redshift range to directly use quasars themselves as cosmological tracers of the matter field.

The clustering of quasars started receiving much attention thanks to the previous programs such as the 2dF Quasar Redshift Survey [2QZ, 189, 190], SDSS-I/II [191, 192, 193] and a combination of quasar samples from the 2QZ and the 2dF-SDSS LRG and quasar Survey [194, 195]. Those studies have revealed that the observed correlation of quasars is the one expected for tracers of the underlying matter distribution and have highlighted the possibility of using quasars to constrain cosmology at higher redshifts than would be possible for galaxy samples to similar magnitude limits. At redshifts $z > 1$, one can use star-forming regions such as Emission Line Galaxies (ELG) which are more prevalent at these high redshifts. Only one recent exploratory measurement has been published using the near-infrared Fiber Multi-Object Spectrograph (FMOS) at the Subaru telescope to detect the Hα lines [FastSound, 143]. The SDSS-IV eBOSS has also acquired data on this tracer in the optical wavelength coverage where the OII] line can be used to measure redshifts in the range $0.6 < z < 1.1$.

The goal of this thesis work is to measure and analyse the clustering of the eBOSS quasar sample between $0.8 \leq z \leq 2.2$ in configuration space to constrain the angular diameter distance $D_{\rm A}(z)$, the Hubble parameter $H(z)$ and the linear growth rate of structure

TABLE 1.2 – Precision of $f\sigma_8$, D_M and H from full-shape analyses using various large-scale surveys. Note that when no values for D_M and H are quoted, the measurement of $f\sigma_8$ has been performed without marginalizing over the AP parameters.

Survey	Reference	z	$f\sigma_8$	D_A	H
2dFGRS	[125]	0.17	13%	–	–
6dFGS	[140]	0.067	13%	–	–
GAMA	[187]	0.18	25%	–	–
		0.38	13.6%	–	–
WiggleZ	[188]	0.44	19.4%	9.4%	9.4%
		0.60	16.2%	6.9%	6.9%
		0.73	16.5%	7.0%	7.2%
VVDS	[146]	0.77	36.7%	–	–
VIPERS	[148]	0.60	21.8%	–	–
		0.86	27.5%	–	–
Subaru FastSound	[143]	1.4	25%	–	–
SDSS MGS	[185]	0.15	35.8%	*	*
SDSS BOSS galaxies	[133]	0.38	9.5%	1.7%	2.7%
		0.51	8.7%	1.6%	2.6 %
		0.61	8.8%	1.8%	2.8%
SDSS BOSS galaxies BAO+FS	[133]	0.38	9.2%	1.5%	2.4%
	[133]	0.51	8.3%	1.4%	2.2%
	[133]	0.61	8.0%	1.4%	2.2%

$f(z)\sigma_8(z)$ at $z_{\rm eff} = 1.52$. This thesis manuscript is structured as follows :

— Chapter 2 presents an overview of the SDSS with a brief description of the technical characteristics of the instruments and a special attention to the eBOSS survey strategy. The observational strategy is defined according to the science requirements : eBOSS aims at measuring the BAO distance at the ∼2% level in redshift $z > 1$ using quasars and ELGs. In order to achieve these requirements, given the instrument capabilities, it sets the requirements for the number of targets that need to be spectroscopically observed. Then, it presents the large-scale structure (LSS) catalogues that have been built from the spectroscopic observations during the first two years of eBOSS. These catalogues represent the fourteenth data release (DR14) of SDSS and the first of eBOSS and we also present how to measure the two-point correlation function of this sample.

— Chapter 3 provides the theoretical concepts that are relevant to model the two-point correlation function in redshift space. Such modeling include three non-linearities : the non-linear evolution of the density and velocity fields, the non-linear relation between dark matter halos and quasars and the non-linear mapping between real and redshift space due to RSD.

— In order to provide unbiased cosmological measurements, we perform a detailed study of modeling and observational systematics. This study is presented in Chapter 4 and also includes the creation of mock catalogues, which are fictive realizations from numerical simulations that reproduce the clustering of the data. These mock catalogues are used as benchmark to test the procedure of analysis.

— The configuration space measurements of this work are presented in Chapter 5 along with the results from several companions studies using the same DR14 quasar sample

and with a comparison of our work compared to previous studies. The final section presents the cosmological implications of the full-shape measurements of this work in the context of the ΛCDM model. We also attempt at providing a discussion on the way to constrain alternative gravity models using our measurements with some constraints for specific models.

The outcome of this thesis is concluded and discussed in Chapter 6, along with several prospects for future work.

Bibliographie

[1] P. J. E. Peebles, *The large-scale structure of the universe* (1980).

[2] E. W. Kolb et M. S. Turner, *Book-Review - the Early Universe*, **80**, 381 (1990).

[3] F. Bernardeau, S. Colombi, E. Gaztañaga, et R. Scoccimarro, *Large-scale structure of the Universe and cosmological perturbation theory*, **367**, 1 (2002), doi:10.1016/S0370-1573(02)00135-7, astro-ph/0112551.

[4] S. Dodelson, *Modern cosmology* (2003).

[5] J. Rich, *Fundamentals of Cosmology* (2010).

[6] S. Serjeant, *Observational Cosmology* (2010).

[7] P. Touboul, G. Métris, M. Rodrigues, Y. André, *et al.*, *MICROSCOPE Mission : First Results of a Space Test of the Equivalence Principle*, Physical Review Letters **119**, 231101 (2017), doi:10.1103/PhysRevLett.119.231101, arXiv:1712.01176 [astro-ph.IM].

[8] A. Einstein, *Kosmologische betrachtungen zur allgemeinen relativitätstheorie*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin), Seite 142-152. (1917).

[9] D. W. Hogg, *Distance measures in cosmology*, ArXiv Astrophysics e-prints (1999), astro-ph/9905116.

[10] V. M. Slipher, in *Publications of the American Astronomical Society* (1922), vol. 4 de *Publications of the American Astronomical Society*, p. 232.

[11] E. Hubble, *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*, Proceedings of the National Academy of Science **15**, 168 (1929), doi:10.1073/pnas.15.3.168.

[12] H. S. Leavitt, *1777 variables in the Magellanic Clouds*, Annals of Harvard College Observatory **60**, 87 (1908).

[13] A. Friedmann, *Über die Krümmung des Raumes*, Zeitschrift fur Physik **10**, 377 (1922), doi:10.1007/BF01332580.

[14] G. Lemaître, *Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques*, Annales de la Société Scientifique de Bruxelles **47**, 49 (1927).

[15] G. Lemaître, *Expansion of the universe, A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulae*, Monthly Notices of the Royal Astronomical Society **91**, 483 (1931).

[16] V. Trimble, *Extragalactic Distance Scales : it H from Hubble (edwin) to Hubble (hubble Telescope)*, **79**, 793 (1997), doi:10.1023/A:1004905101332.

[17] H. Bondi et T. Gold, *The Steady-State Theory of the Expanding Universe*, **108**, 252 (1948), doi:10.1093/mnras/108.3.252.

[18] F. Hoyle, *A New Model for the Expanding Universe*, **108**, 372 (1948), doi:10.1093/mnras/108.5.372.

[19] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, *et al.*, *Planck 2015 results. XIII. Cosmological parameters*, **594**, A13 (2016), doi:10.1051/0004-6361/201525830, arXiv:1502.01589.

[20] G. Gamow, *Expanding Universe and the Origin of Elements*, Physical Review **70**, 572 (1946), doi:10.1103/PhysRev.70.572.2.

[21] G. Gamow, *The Evolution of the Universe*, **162**, 680 (1948), doi:10.1038/162680a0.

[22] R. A. Alpher et R. Herman, *Evolution of the Universe*, Nature **162**, 774 (1948), doi:10.1038/162774b0.

[23] A. A. Penzias et R. W. Wilson, *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*, The Astrophysical Journal **142**, 419 (1965), doi:10.1086/148307.

[24] A. A. Penzias et R. W. Wilson, *Measurement of the Flux Density of CAS a at 4080 Mc/s.*, The Astrophysical Journal **142**, 1149 (1965), doi:10.1086/148384.

[25] R. H. Dicke, P. J. E. Peebles, P. G. Roll, et D. T. Wilkinson, *Cosmic Black-Body Radiation.*, **142**, 414 (1965), doi:10.1086/148306.

[26] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, *et al.*, *Structure in the COBE differential microwave radiometer first-year maps*, **396**, L1 (1992), doi:10.1086/186504.

[27] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, *et al.*, *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Final Maps and Results*, **208**, 20 (2013), doi:10.1088/0067-0049/208/2/20, arXiv:1212.5225.

[28] M. White, éd., *Anisotropies in the CMB* (1999), astro-ph/9903232.

[29] J. Silk, *Cosmic Black-Body Radiation and Galaxy Formation*, **151**, 459 (1968), doi:10.1086/149449.

[30] A. H. Guth, *Inflationary universe : A possible solution to the horizon and flatness problems*, **23**, 347 (1981), doi:10.1103/PhysRevD.23.347.

[31] A. R. Liddle et D. H. Lyth, *Cosmological Inflation and Large-Scale Structure* (2000).

[32] A. D. Linde, *A new inflationary universe scenario : A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems*, Physics Letters B **108**, 389 (1982), doi:10.1016/0370-2693(82)91219-9.

[33] A. Albrecht et P. J. Steinhardt, *Cosmology for grand unified theories with radiatively induced symmetry breaking*, Physical Review Letters **48**, 1220 (1982), doi:10.1103/PhysRevLett.48.1220.

[34] B. A. Reid et M. White, *Towards an accurate model of the redshift-space clustering of haloes in the quasi-linear regime*, **417**, 1913 (2011), doi:10.1111/j.1365-2966.2011.19379.x, arXiv:1105.4165.

[35] U. Seljak et M. Zaldarriaga, *A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies*, **469**, 437 (1996), doi:10.1086/177793, astro-ph/9603033.

[36] A. Lewis, A. Challinor, et A. Lasenby, *Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models*, **538**, 473 (2000), doi:10.1086/309179, astro-ph/9911177.

[37] J. Lesgourgues, *The Cosmic Linear Anisotropy Solving System (CLASS) I : Overview*, ArXiv e-prints (2011), arXiv:1104.2932 [astro-ph.IM].

[38] S. D. M. White et M. J. Rees, *Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering*, **183**, 341 (1978), doi:10.1093/mnras/183.3.341.

[39] J. Baur, N. Palanque-Delabrouille, C. Yèche, C. Magneville, *et al.*, *Lyman-alpha forests cool warm dark matter*, **8**, 012 (2016), doi:10.1088/1475-7516/2016/08/012, arXiv:1512.01981.

[40] M. Tegmark, M. R. Blanton, M. A. Strauss, F. Hoyle, *et al.*, *The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey*, **606**, 702 (2004), doi:10.1086/382125, astro-ph/0310725.

[41] J. H. Oort, *Note on the distribution of luminosities of K and M giants*, **6**, 289 (1932).

[42] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, Helvetica Physica Acta **6**, 110 (1933).

[43] V. C. Rubin et J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, The Astrophysical Journal **159**, 379 (1970), doi:10.1086/150317.

[44] E. Corbelli, S. Lorenzoni, R. Walterbos, R. Braun, *et al.*, *A wide-field H I mosaic of Messier 31. II. The disk warp, rotation, and the dark matter halo*, **511**, A89 (2010), doi:10.1051/0004-6361/200913297, arXiv:0912.4133.

[45] L. L. Watkins, N. W. Evans, et J. H. An, *The masses of the Milky Way and Andromeda galaxies*, **406**, 264 (2010), doi:10.1111/j.1365-2966.2010.16708.x, arXiv:1002.4565.

[46] R. D. Peccei et H. R. Quinn, *Constraints imposed by CP conservation in the presence of pseudoparticles*, **16**, 1791 (1977), doi:10.1103/PhysRevD.16.1791.

[47] M. Milgrom, *A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis*, **270**, 365 (1983), doi:10.1086/161130.

[48] S. J. Maddox, G. Efstathiou, W. J. Sutherland, et J. Loveday, *Galaxy correlations on large scales*, Monthly Notices of the Royal Astronomical Society **242**, 43P (1990), doi:10.1093/mnras/242.1.43P.

[49] S. D. M. White, J. F. Navarro, A. E. Evrard, et C. S. Frenk, *The baryon content of galaxy clusters : a challenge to cosmological orthodoxy*, Nature **366**, 429 (1993), doi:10.1038/366429a0.

[50] E. D. Loh et E. J. Spillar, *A measurement of the mass density of the universe*, Astrophysical Journal **307**, L1 (1986), doi:10.1086/184717.

[51] A. Nusser et A. Dekel, *Omega and the initial fluctuations from velocity and density fields*, Astrophysical Journal **405**, 437 (1993), doi:10.1086/172376.

[52] G. Efstathiou, W. J. Sutherland, et S. J. Maddox, *The cosmological constant and cold dark matter*, Nature **348**, 705 (1990), doi:10.1038/348705a0.

[53] L. M. Krauss et M. S. Turner, *The cosmological constant is back*, General Relativity and Gravitation **27**, 1137 (1995), doi:10.1007/BF02108229, astro-ph/9504003.

[54] M. J. Pierce, D. L. Welch, R. D. McClure, S. van den Bergh, *et al.*, *The Hubble constant and Virgo cluster distance from observations of Cepheid variables*, **371**, 385 (1994), doi:10.1038/371385a0.

[55] W. L. Freedman, B. F. Madore, J. R. Mould, R. Hill, *et al.*, *Distance to the Virgo cluster galaxy M100 from Hubble Space Telescope observations of Cepheids*, **371**, 757 (1994), doi:10.1038/371757a0.

[56] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, *et al.*, *Measurements of* Ω *and* Λ *from 42 High-Redshift Supernovae*, The Astrophysical Journal **517**, 565 (1999), doi:10.1086/307221, arXiv:9812133 [astro-ph].

[57] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, *et al.*, *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, The Astronomical Journal **116**, 1009 (1998), doi:10.1086/300499.

[58] A. G. Riess, W. H. Press, et R. P. Kirshner, *A Precise Distance Indicator : Type IA Supernova Multicolor Light-Curve Shapes*, **473**, 88 (1996), doi:10.1086/178129, astro-ph/9604143.

[59] A. Sakharov, *Vacuum quantum fluctuations in curved space and the theory of gravitation*, General Relativity and Gravitation **32**, 365 (2000).

[60] S. Weinberg, *The cosmological constant problem*, Reviews of Modern Physics **61**, 1 (1989), doi:10.1103/RevModPhys.61.1.

[61] J. Martin, *Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask)*, Comptes Rendus Physique p. 89 (2012), arXiv:1205.3365.

[62] B. Ratra et P. J. E. Peebles, *Cosmological consequences of a rolling homogeneous scalar field*, **37**, 3406 (1988), doi:10.1103/PhysRevD.37.3406.

[63] P. G. Ferreira et M. Joyce, *Structure Formation with a Self-Tuning Scalar Field*, Physical Review Letters **79**, 4740 (1997), doi:10.1103/PhysRevLett.79.4740, astro-ph/9707286.

[64] I. Zlatev, L. Wang, et P. J. Steinhardt, *Quintessence, Cosmic Coincidence, and the Cosmological Constant*, Physical Review Letters **82**, 896 (1999), doi:10.1103/PhysRevLett.82.896, astro-ph/9807002.

[65] M. Chevallier et D. Polarski, *Accelerating Universes with Scaling Dark Matter*, International Journal of Modern Physics D **10**, 213 (2001), doi:10.1142/S0218271801000822, gr-qc/0009008.

[66] E. V. Linder, *Exploring the Expansion History of the Universe*, Physical Review Letters **90**, 091301 (2003), doi:10.1103/PhysRevLett.90.091301, astro-ph/0208512.

[67] S. M. Carroll, V. Duvvuri, M. Trodden, et M. S. Turner, *Is cosmic speed-up due to new gravitational physics ?*, **70**, 043528 (2004), doi:10.1103/PhysRevD.70.043528, astro-ph/0306438.

[68] C. P. L. Berry et J. R. Gair, *Erratum : Linearized f(R) gravity : gravitational radiation and solar system tests [Phys. Rev. D 83, 104022 (2011)]*, **85**, 089906 (2012), doi:10.1103/PhysRevD.85.089906.

[69] J. Noller, F. von Braun-Bates, et P. G. Ferreira, *Relativistic scalar fields and the quasistatic approximation in theories of modified gravity*, **89**, 023521 (2014), doi:10.1103/PhysRevD.89.023521, arXiv:1310.3266 [astro-ph.CO].

[70] E. Jennings, C. M. Baugh, B. Li, G.-B. Zhao, *et al.*, *Redshift-space distortions in f(R) gravity*, **425**, 2128 (2012), doi:10.1111/j.1365-2966.2012.21567.x, arXiv:1205.2698 [astro-ph.CO].

[71] J. Khoury et A. Weltman, *Chameleon cosmology*, **69**, 044026 (2004), doi:10.1103/PhysRevD.69.044026, astro-ph/0309411.

[72] D. F. Mota et D. J. Shaw, *Evading equivalence principle violations, cosmological, and other experimental constraints in scalar field theories with a strong coupling to matter*, **75**, 063501 (2007), doi:10.1103/PhysRevD.75.063501, hep-ph/0608078.

[73] B. Li et J. D. Barrow, *Cosmology of f(R) gravity in the metric variational approach*, **75**, 084010 (2007), doi:10.1103/PhysRevD.75.084010, gr-qc/0701111.

[74] P. Brax, C. van de Bruck, A.-C. Davis, et D. J. Shaw, *f(R) gravity and chameleon theories*, **78**, 104021 (2008), doi:10.1103/PhysRevD.78.104021, arXiv:0806.3415.

[75] I. Navarro et K. Van Acoleyen, *f(R) actions, cosmic acceleration and local tests of gravity*, **2**, 022 (2007), doi:10.1088/1475-7516/2007/02/022, gr-qc/0611127.

[76] W. Hu et I. Sawicki, *Models of f(R) cosmic acceleration that evade solar system tests*, **76**, 064004 (2007), doi:10.1103/PhysRevD.76.064004, arXiv:0705.1158.

[77] G. Dvali, G. Gabadadze, et M. Porrati, *4D gravity on a brane in 5D Minkowski space*, Physics Letters B **485**, 208 (2000), doi:10.1016/S0370-2693(00)00669-9, arXiv:0005016 [hep-th].

[78] A. Vainshtein, *To the problem of nonvanishing gravitation mass*, Physics Letters B **39**, 3 (1972).

[79] C. Deffayet, G. Dvali, G. Gabadadze, et A. Vainshtein, *Nonperturbative continuity in graviton mass versus perturbative discontinuity*, Physical Review D **65**, 044026 (2002), doi:10.1103/PhysRevD.65.044026, arXiv:0106001 [hep-th].

[80] K. Koyama et R. Maartens, *Structure formation in the Dvali–Gabadadze–Porrati cosmological model*, J. Cosmol. Astropart. Phys. **2006**, 016 (2006), doi:10.1088/1475-7516/2006/01/016, astro-ph/0511634.

[81] M. A. Luty, M. Porrati, et R. Rattazzi, *Strong interactions and stability in the dgp model*, Journal of High Energy Physics **2003**, 029 (2003).

[82] W. Fang, S. Wang, W. Hu, Z. Haiman, *et al.*, *Challenges to the DGP model from horizon-scale growth and geometry*, Physical Review D **78**, 103509 (2008), doi:10.1103/PhysRevD.78.103509, arXiv:0808.2208.

[83] H. van Dam et M. Veltman, *Massive and massless Yang-Mills and gravitational fields*, Nuclear Physics B **22**, 397 (1970), doi:10.1016/0550-3213(70)90416-5.

[84] V. I. Zakharov, *Linearized Gravitation Theory and the Graviton Mass*, ZhETF Pisma Redaktsiiu **12**, 447 (1970).

[85] D. Boulware et S. Deser, *Inconsistency of finite range gravitation*, Phys.Lett. **B40**, 227 (1972), doi:10.1016/0370-2693(72)90418-2.

[86] C. de Rham et G. Gabadadze, *Generalization of the Fierz-Pauli action*, Physical Review D **82**, 044020 (2010), doi:10.1103/PhysRevD.82.044020, arXiv:1007.0443.

[87] C. de Rham, G. Gabadadze, et A. J. Tolley, *Resummation of Massive Gravity*, Physical Review Letters **106**, 231101 (2011), doi:10.1103/PhysRevLett.106.231101, arXiv:1011.1232.

[88] A. De Felice, A. E. Gümrükçüoğlu, et S. Mukohyama, *Massive Gravity : Nonlinear Instability of a Homogeneous and Isotropic Universe*, Physical Review Letters **109**, 171101 (2012), doi:10.1103/PhysRevLett.109.171101, arXiv:1206.2080.

[89] A. E. Gümrükçüoğlu, C. Lin, et S. Mukohyama, *Anisotropic Friedmann–Robertson–Walker universe from nonlinear massive gravity*, Physics Letters B **717**, 295 (2012), doi:10.1016/j.physletb.2012.09.049, arXiv:1206.2723.

[90] A. Nicolis, R. Rattazzi, et E. Trincherini, *Galileon as a local modification of gravity*, Physical Review D **79**, 064036 (2009), doi:10.1103/PhysRevD.79.064036, arXiv:0811.2197.

[91] E. Babichev et G. Esposito-Farèse, *Time-dependent spherically symmetric covariant Galileons*, Physical Review D **87**, 044032 (2013), doi:10.1103/PhysRevD.87.044032, arXiv:1212.1394.

[92] J. Neveu, V. Ruhlmann-Kleider, P. Astier, M. Besançon, *et al.*, *Constraining the ΛCDM and Galileon models with recent cosmological data*, **600**, A40 (2017), doi:10.1051/0004-6361/201628878, arXiv:1605.02627 [gr-qc].

[93] J. G. Williams, S. G. Turyshev, et D. H. Boggs, *Lunar laser ranging tests of the equivalence principle*, Classical and Quantum Gravity **29**, 184004 (2012), doi:10.1088/0264-9381/29/18/184004, arXiv:1203.2150 [gr-qc].

[94] A. Barreira, B. Li, W. A. Hellwing, C. M. Baugh, *et al.*, *Nonlinear structure formation in the cubic Galileon gravity model*, Journal of Cosmology and Astroparticle Physics **2013**, 027 (2013), doi:10.1088/1475-7516/2013/10/027, arXiv:1306.3219.

[95] A. Barreira, B. Li, C. M. Baugh, et S. Pascoli, *Spherical collapse in Galileon gravity : fifth force solutions, halo mass function and halo bias*, Journal of Cosmology and Astroparticle Physics **2013**, 056 (2013), doi:10.1088/1475-7516/2013/11/056, arXiv:1308.3699.

[96] J. Neveu, V. Ruhlmann-Kleider, A. Conley, N. Palanque-Delabrouille, *et al.*, *Experimental constraints on the uncoupled Galileon model from SNLS3 data and other cosmological probes*, **555**, A53 (2013), doi:10.1051/0004-6361/201321256, arXiv:1302.2786 [gr-qc].

[97] B. Li, A. Barreira, C. M. Baugh, W. A. Hellwing, *et al.*, *Simulating the quartic Galileon gravity model on adaptively refined meshes*, Journal of Cosmology and Astroparticle Physics **2013**, 012 (2013), doi:10.1088/1475-7516/2013/11/012, 1308.3491.

[98] J. Neveu, V. Ruhlmann-Kleider, P. Astier, M. Besançon, *et al.*, *First experimental constraints on the disformally coupled Galileon model*, **569**, A90 (2014), doi:10.1051/0004-6361/201423758, arXiv:1403.0854 [gr-qc].

[99] P. J. E. Peebles, *Statistical Analysis of Catalogs of Extragalactic Objects. I. Theory*, **185**, 413 (1973), doi:10.1086/152431.

[100] S. A. Gregory et L. A. Thompson, *Low mass portion of the galaxy clustering spectrum*, **274**, 450 (1978), doi:10.1038/274450a0.

[101] S. A. Gregory, L. A. Thompson, et W. G. Tifft, *The Perseus supercluster*, **243**, 411 (1981), doi:10.1086/158608.

[102] R. P. Kirshner, A. Oemler, Jr., P. L. Schechter, et S. A. Shectman, *A million cubic megaparsec void in Bootes*, **248**, L57 (1981), doi:10.1086/183623.

[103] I. B. Zeldovich, J. Einasto, et S. F. Shandarin, *Giant voids in the universe*, **300**, 407 (1982), doi:10.1038/300407a0.

[104] J. Huchra, M. Davis, D. Latham, et J. Tonry, *A survey of galaxy redshifts. IV - The data*, **52**, 89 (1983), doi:10.1086/190860.

[105] V. de Lapparent, M. J. Geller, et J. P. Huchra, *A Slice of the Universe*, **302**, L1 (1986), doi:10.1086/184625.

[106] M. J. Geller et J. P. Huchra, *Mapping the Universe*, Science **246**, 897 (1989), doi:10.1126/science.246.4932.897.

[107] M. S. Vogeley, C. Park, M. J. Geller, et J. P. Huchra, *Large-Scale Clustering of Galaxies in the CfA Redshift Survey*, **391**, L5 (1992), doi:10.1086/186385.

[108] G. Efstathiou, N. Kaiser, W. Saunders, A. Lawrence, *et al.*, *Large-scale clustering of IRAS galaxies.*, **247**, 10P (1990).

[109] W. Saunders, C. Frenk, M. Rowan-Robinson, G. Efstathiou, *et al.*, *The density field of the local Universe*, **349**, 32 (1991), doi:10.1038/349032a0.

[110] K. B. Fisher, M. Davis, M. A. Strauss, A. Yahil, *et al.*, *Clustering in the 1.2-Jy IRAS galaxy redshift survey - II. Redshift distortions and dzeta(rp, pi).*, **267**, 927 (1994), doi:10.1093/mnras/267.4.927.

[111] K. B. Fisher, J. P. Huchra, M. A. Strauss, M. Davis, *et al.*, *The IRAS 1.2 Jy Survey : Redshift Data*, The Astrophysical Journal Supplement Series **100**, 69 (1995), doi:10.1086/192208.

[112] A. F. Heavens et A. N. Taylor, *A spherical harmonic analysis of redshift space*, **275**, 483 (1995), doi:10.1093/mnras/275.2.483.

[113] W. E. Ballinger, A. F. Heavens, et A. N. Taylor, *The real-space power spectrum of IRAS galaxies on large scales and the redshift distortion*, **276**, L59 (1995), doi:10.1093/mnras/276.1.L59.

[114] S. Cole, K. B. Fisher, et D. H. Weinberg, *Fourier Analysis of Redshift Space Distortions and the Determination of Omega*, **267**, 785 (1994), doi:10.1093/mnras/267.3.785.

[115] W. Saunders, W. J. Sutherland, S. J. Maddox, O. Keeble, *et al.*, *The PSCz catalogue*, **317**, 55 (2000), doi:10.1046/j.1365-8711.2000.03528.x.

[116] H. Tadros, W. E. Ballinger, A. N. Taylor, A. F. Heavens, *et al.*, *Spherical harmonic analysis of the PSCz galaxy catalogue : redshift distortions and the real-space power spectrum*, **305**, 527 (1999), doi:10.1046/j.1365-8711.1999.02409.x.

[117] M. Tegmark, A. J. S. Hamilton, et Y. Xu, *The power spectrum of galaxies in the 2dF 100k redshift survey*, **335**, 887 (2002), doi:10.1046/j.1365-8711.2002.05622.x.

[118] A. Albrecht, G. Bernstein, R. Cahn, W. L. Freedman, *et al.*, *Report of the Dark Energy Task Force*, ArXiv e-prints astro-ph/0609591 (2006).

[119] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, *et al.*, *Observational probes of cosmic acceleration*, **530**, 87 (2013), doi:10.1016/j.physrep.2013.05.001, arXiv:1201.2434.

[120] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, *et al.*, *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies*, **633**, 560 (2005), doi:10.1086/466512.

[121] W. J. Percival, B. A. Reid, D. J. Eisenstein, N. A. Bahcall, *et al.*, *Baryon acoustic oscillations in the Sloan Digital Sky Survey Data Release 7 galaxy sample*, **401**, 2148 (2010), doi:10.1111/j.1365-2966.2009.15812.x, arXiv:0907.1660.

[122] N. Padmanabhan, X. Xu, D. J. Eisenstein, R. Scalzo, *et al.*, *A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - I. Methods and application to the Sloan Digital Sky Survey*, **427**, 2132 (2012), doi:10.1111/j.1365-2966.2012.21888.x, arXiv:1202.0090.

[123] W. J. Percival, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, *et al.*, *The 2dF Galaxy Redshift Survey : the power spectrum and the matter content of the Universe*, **327**, 1297 (2001), doi:10.1046/j.1365-8711.2001.04827.x.

[124] S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, *et al.*, *The 2dF Galaxy Redshift Survey : power-spectrum analysis of the final data set and cosmological implications*, **362**, 505 (2005), doi:10.1111/j.1365-2966.2005.09318.x.

[125] W. J. Percival, D. Burkey, A. Heavens, A. Taylor, *et al.*, *The 2dF Galaxy Redshift Survey : spherical harmonics analysis of fluctuations in the final catalogue*, **353**, 1201 (2004), doi:10.1111/j.1365-2966.2004.08146.x.

[126] C. Heymans, L. Van Waerbeke, L. Miller, T. Erben, *et al.*, *CFHTLenS : the Canada-France-Hawaii Telescope Lensing Survey*, **427**, 146 (2012), doi:10.1111/j.1365-2966.2012.21952.x, arXiv:1210.0032.

[127] W. M. Wood-Vasey, G. Miknaitis, C. W. Stubbs, S. Jha, *et al.*, *Observational Constraints on the Nature of Dark Energy : First Cosmological Results from the ESSENCE Supernova Survey*, **666**, 694 (2007), doi:10.1086/518642.

[128] J. A. Frieman, B. Bassett, A. Becker, C. Choi, *et al.*, *The Sloan Digital Sky Survey-II Supernova Survey : Technical Summary*, **135**, 338 (2008), doi:10.1088/0004-6256/135/1/338, arXiv:0708.2749.

[129] W. L. Freedman, C. R. Burns, M. M. Phillips, P. Wyatt, *et al.*, *The Carnegie Supernova Project : First Near-Infrared Hubble Diagram to z 0.7*, **704**, 1036 (2009), doi:10.1088/0004-637X/704/2/1036, arXiv:0907.4524.

[130] A. G. Riess, L.-G. Strolger, S. Casertano, H. C. Ferguson, *et al.*, *New Hubble Space Telescope Discoveries of Type Ia Supernovae at z >= 1 : Narrowing Constraints on the Early Behavior of Dark Energy*, **659**, 98 (2007), doi:10.1086/510378.

[131] D. J. Eisenstein, D. H. Weinberg, E. Agol, H. Aihara, *et al.*, *SDSS-III : Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems*, **142**, 72 (2011), doi:10.1088/0004-6256/142/3/72, arXiv:1101.1529.

[132] K. S. Dawson, D. J. Schlegel, C. P. Ahn, S. F. Anderson, *et al.*, *The Baryon Oscillation Spectroscopic Survey of SDSS-III*, **145**, 10 (2013), doi:10.1088/0004-6256/145/1/10, arXiv:1208.0022.

[133] S. Alam, M. Ata, S. Bailey, F. Beutler, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : cosmological analysis of the DR12 galaxy sample*, **470**, 2617 (2017), doi:10.1093/mnras/stx721, arXiv:1607.03155.

[134] M. R. Blanton, H. Lin, R. H. Lupton, F. M. Maley, *et al.*, *An Efficient Targeting Strategy for Multiobject Spectrograph Surveys : the Sloan Digital Sky Survey "Tiling" Algorithm*, **125**, 2276 (2003), doi:10.1086/344761, astro-ph/0105535.

[135] K. S. Dawson, J.-P. Kneib, W. J. Percival, S. Alam, *et al.*, *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey : Overview and Early Data*, **151**, 44 (2016), doi:10.3847/0004-6256/151/2/44, arXiv:1508.04473.

[136] The Dark Energy Survey Collaboration, *The Dark Energy Survey*, ArXiv Astrophysics e-prints (2005), astro-ph/0510346.

[137] DES Collaboration, T. M. C. Abbott, F. B. Abdalla, A. Alarcon, *et al.*, *Dark Energy Survey Year 1 Results : Cosmological Constraints from Galaxy Clustering and Weak Lensing*, ArXiv e-prints (2017), arXiv:1708.01530.

[138] S. P. Driver, P. Norberg, I. K. Baldry, S. P. Bamford, *et al.*, *GAMA : towards a physical understanding of galaxy formation*, Astronomy and Geophysics **50**, 5.12 (2009), doi:10.1111/j.1468-4004.2009.50512.x, arXiv:0910.5123.

[139] F. Beutler, C. Blake, M. Colless, D. H. Jones, *et al.*, *The 6dF Galaxy Survey : baryon acoustic oscillations and the local Hubble constant*, **416**, 3017 (2011), doi:10.1111/j.1365-2966.2011.19250.x, arXiv:1106.3366.

[140] F. Beutler, C. Blake, M. Colless, D. H. Jones, *et al.*, *The 6dF Galaxy Survey : z≈ 0 measurements of the growth rate and σ₈*, **423**, 3430 (2012), doi:10.1111/j.1365-2966.2012.21136.x, arXiv:1204.4725.

[141] C. Blake, E. A. Kazin, F. Beutler, T. M. Davis, *et al.*, *The WiggleZ Dark Energy Survey : mapping the distance-redshift relation with baryon acoustic oscillations*, **418**, 1707 (2011), doi:10.1111/j.1365-2966.2011.19592.x, arXiv:1108.2635.

[142] E. A. Kazin, J. Koda, C. Blake, N. Padmanabhan, *et al.*, *The WiggleZ Dark Energy Survey : improved distance measurements to z = 1 with reconstruction of the baryonic acoustic feature*, **441**, 3524 (2014), doi:10.1093/mnras/stu778, arXiv:1401.0358.

[143] T. Okumura, C. Hikage, T. Totani, M. Tonegawa, *et al.*, *The Subaru FMOS galaxy redshift survey (FastSound). IV. New constraint on gravity theory from redshift space distortions at z*
1.4, **68**, 38 (2016), doi:10.1093/pasj/psw029, arXiv:1511.08083.

[144] S. Miyazaki, Y. Komiyama, H. Nakaya, Y. Kamata, *et al.*, in *Ground-based and Airborne Instrumentation for Astronomy IV* (2012), vol. 8446, p. 84460Z.

[145] S. J. Lilly, O. Le Fèvre, A. Renzini, G. Zamorani, *et al.*, *zCOSMOS : A Large VLT/VIMOS Redshift Survey Covering 0 < z < 3 in the COSMOS Field*, The Astrophysical Journal Supplement Series **172**, 70 (2007), doi:10.1086/516589.

[146] L. Guzzo, M. Pierleoni, B. Meneux, E. Branchini, *et al.*, *A test of the nature of cosmic acceleration using galaxy redshift distortions*, **451**, 541 (2008), doi:10.1038/nature06555, arXiv:0802.1944.

[147] O. Le Fèvre, P. Cassata, O. Cucciati, B. Garilli, *et al.*, *The VIMOS VLT Deep Survey final data release : a spectroscopic sample of 35 016 galaxies and AGN out to z 6.7 selected with 17.5 ≤ i_{AB} ≤ 24.75*, **559**, A14 (2013), doi:10.1051/0004-6361/201322179, arXiv:1307.0545.

[148] A. Pezzotta, S. de la Torre, J. Bel, B. R. Granett, *et al.*, *The VIMOS Public Extragalactic Redshift Survey (VIPERS). The growth of structure at 0.5 < z < 1.2 from redshift-space distortions in the clustering of the PDR-2 final sample*, **604**, A33 (2017), doi:10.1051/0004-6361/201630295.

[149] F. G. Mohammad, S. de la Torre, D. Bianchi, L. Guzzo, *et al.*, *Group-galaxy correlations in redshift space as a probe of the growth of structure*, **458**, 1948 (2016), doi:10.1093/mnras/stw411.

[150] G. J. Hill, K. Gebhardt, E. Komatsu, N. Drory, *et al.*, in *Panoramic Views of Galaxy Formation and Evolution* (2008), vol. 399, p. 115, arXiv:0806.0183.

[151] K. C. Chambers, E. A. Magnier, N. Metcalfe, H. A. Flewelling, *et al.*, *The Pan-STARRS1 Surveys*, ArXiv e-prints arXiv :1612.05560 (2016), arXiv:1612.05560.

[152] Gaia Collaboration, T. Prusti, J. H. J. de Bruijne, A. G. A. Brown, *et al.*, *The Gaia mission*, **595**, A1 (2016), doi:10.1051/0004-6361/201629272.

[153] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, *et al.*, *The DESI Experiment Part I : Science,Targeting, and Survey Design*, ArXiv e-prints (2016), arXiv:1611.00036 [astro-ph.IM].

[154] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, *et al.*, *The DESI Experiment Part II : Instrument Design*, ArXiv e-prints (2016), arXiv:1611.00037 [astro-ph.IM].

[155] M. Takada, R. S. Ellis, M. Chiba, J. E. Greene, *et al.*, *Extragalactic science, cosmology, and Galactic archaeology with the Subaru Prime Focus Spectrograph*, Publications of the Astronomical Society of Japan **66**, R1 (2014), doi:10.1093/pasj/pst019, arXiv:1206.0737.

[156] L. Amendola, S. Appleby, D. Bacon, T. Baker, *et al.*, *Cosmology and Fundamental Physics with the Euclid Satellite*, Living Reviews in Relativity **16**, 6 (2013), doi:10.12942/lrr-2013-6, arXiv:1206.1225.

[157] LSST Dark Energy Science Collaboration, *Large Synoptic Survey Telescope : Dark Energy Science Collaboration*, ArXiv e-prints (2012), arXiv:1211.0310 [astro-ph.CO].

[158] R. S. de Jong, S. Barden, O. Bellido-Tirado, J. Brynnel, *et al.*, in *Ground-based and Airborne Instrumentation for Astronomy V* (2014), vol. 9147 de , p. 91470M.

[159] D. Spergel, N. Gehrels, J. Breckinridge, M. Donahue, *et al.*, *Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA Final Report*, ArXiv e-prints arXiv :1305.5422 (2013), arXiv:1305.5422.

[160] R. Maartens, F. B. Abdalla, M. Jarvis, M. G. Santos, *et al.*, *Cosmology with the SKA – overview*, ArXiv e-prints arXiv :1501.04076 (2015), arXiv:1501.04076.

[161] S. Camera, M. G. Santos, et R. Maartens, *Probing primordial non-Gaussianity with SKA galaxy redshift surveys : a fully relativistic analysis*, **448**, 1035 (2015), doi:10.1093/mnras/stv040, arXiv:1409.8286.

[162] C. Alcock et B. Paczynski, *An evolution free test for non-zero cosmological constant*, **281**, 358 (1979), doi:10.1038/281358a0.

[163] J. A. Peacock, *The evolution of clustering and bias in the galaxy distribution*, Philosophical Transactions of the Royal Society of London Series A **357**, 133 (1999), doi:10.1098/rsta.1999.0319.

[164] D. J. Eisenstein, H.-J. Seo, E. Sirko, et D. N. Spergel, *Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak*, **664**, 675 (2007), doi:10.1086/518712, astro-ph/0604362.

[165] M. Kamionkowski, B. Ratra, D. N. Spergel, et N. Sugiyama, *Cosmic Background Radiation Anisotropy in an Open Inflation, Cold Dark Matter Cosmogony*, **434**, L1 (1994), doi:10.1086/187558.

[166] D. J. Eisenstein et W. Hu, *Baryonic Features in the Matter Transfer Function*, **496**, 605 (1998), doi:10.1086/305424.

[167] D. J. Eisenstein, J. Annis, J. E. Gunn, A. S. Szalay, *et al.*, *Spectroscopic Target Selection for the Sloan Digital Sky Survey : The Luminous Red Galaxy Sample*, **122**, 2267 (2001), doi:10.1086/323717.

[168] H.-J. Seo et D. J. Eisenstein, *Baryonic Acoustic Oscillations in Simulated Galaxy Redshift Surveys*, **633**, 575 (2005), doi:10.1086/491599, astro-ph/0507338.

[169] D. J. Eisenstein, H.-J. Seo, et M. White, *On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter*, **664**, 660 (2007), doi:10.1086/518755, astro-ph/0604361.

[170] W. Hu et Z. Haiman, *Redshifting rings of power*, **68**, 063004 (2003), doi:10.1103/PhysRevD.68.063004, astro-ph/0306053.

[171] L. Samushia, B. A. Reid, M. White, W. J. Percival, *et al.*, *The clustering of galaxies in the SDSS-III DR9 Baryon Oscillation Spectroscopic Survey : testing deviations from Λ and general relativity using anisotropic clustering of galaxies*, **429**, 1514 (2013), doi:10.1093/mnras/sts443, arXiv:1206.5309.

[172] N. Kaiser, *Clustering in real space and in redshift space*, **227**, 1 (1987).

[173] J. C. Jackson, *A critique of Rees's theory of primordial gravitational radiation*, **156**, 1P (1972), doi:10.1093/mnras/156.1.1P, arXiv:0810.3908.

[174] A. J. S. Hamilton, in *The Evolving Universe*, édité par D. Hamilton (1998), vol. 231 de *Astrophysics and Space Science Library*, p. 185, astro-ph/9708102.

[175] Y. B. Zel'dovich, *Gravitational instability : An approximate theory for large density perturbations.*, **5**, 84 (1970).

[176] J. A. Peacock, S. Cole, P. Norberg, C. M. Baugh, *et al.*, *A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey*, **410**, 169 (2001), astro-ph/0103143.

[177] W. J. Percival et M. White, *Testing cosmological structure formation using redshift-space distortions*, **393**, 297 (2009), doi:10.1111/j.1365-2966.2008.14211.x, arXiv:0808.0003.

[178] O. Le Fèvre, G. Vettolani, B. Garilli, L. Tresse, *et al.*, *The VIMOS VLT deep survey. First epoch VVDS-deep survey : 11 564 spectra with $17.5 \leq i_{AB} \leq 24$, and the redshift distribution over $0 \leq z \leq 5$*, **439**, 845 (2005), doi:10.1051/0004-6361:20041960, astro-ph/0409133.

[179] B. A. Reid, L. Samushia, M. White, W. J. Percival, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : measurements of the growth of structure and expansion rate at $z = 0.57$ from anisotropic clustering*, **426**, 2719 (2012), doi:10.1111/j.1365-2966.2012.21779.x, arXiv:1203.6641.

[180] T. Delubac, J. E. Bautista, N. G. Busca, J. Rich, *et al.*, *Baryon acoustic oscillations in the Lyα forest of BOSS DR11 quasars*, **574**, A59 (2015), doi:10.1051/0004-6361/201423969, arXiv:1404.1801.

[181] A. Font-Ribera, D. Kirkby, N. Busca, J. Miralda-Escudé, *et al.*, *Quasar-Lyman α forest cross-correlation from BOSS DR11 : Baryon Acoustic Oscillations*, **5**, 027 (2014), doi:10.1088/1475-7516/2014/05/027, arXiv:1311.1767.

[182] J. E. Bautista, N. G. Busca, J. Guy, J. Rich, *et al.*, *Measurement of baryon acoustic oscillation correlations at $z = 2.3$ with SDSS DR12 Lyα-Forests*, **603**, A12 (2017), doi:10.1051/0004-6361/201730533, arXiv:1702.00176.

[183] H. du Mas des Bourboux, J.-M. Le Goff, M. Blomqvist, N. G. Busca, *et al.*, *Baryon acoustic oscillations from the complete SDSS-III Lyα-quasar cross-correlation function at $z = 2.4$*, ArXiv e-prints (2017), arXiv:1708.02225.

[184] A. J. Ross, W. J. Percival, et M. Manera, *The information content of anisotropic Baryon Acoustic Oscillation scale measurements*, **451**, 1331 (2015), doi:10.1093/mnras/stv966, arXiv:1501.05571.

[185] C. Howlett, A. J. Ross, L. Samushia, W. J. Percival, *et al.*, *The clustering of the SDSS main galaxy sample - II. Mock galaxy catalogues and a measurement of the growth of structure from redshift space distortions at $z = 0.15$*, **449**, 848 (2015), doi:10.1093/mnras/stu2693, arXiv:1409.3238.

[186] F. Beutler, C. Blake, J. Koda, F. A. Marín, *et al.*, *The BOSS-WiggleZ overlap region - I. Baryon acoustic oscillations*, **455**, 3230 (2016), doi:10.1093/mnras/stv1943, arXiv:1506.03900.

[187] C. Blake, I. K. Baldry, J. Bland-Hawthorn, L. Christodoulou, *et al.*, *Galaxy And Mass Assembly (GAMA) : improved cosmic growth measurements using multiple tracers of large-scale structure*, **436**, 3089 (2013), doi:10.1093/mnras/stt1791, arXiv:1309.5556.

[188] C. Blake, S. Brough, M. Colless, C. Contreras, *et al.*, *The WiggleZ Dark Energy Survey : joint measurements of the expansion and growth history at z 1*, **425**, 405 (2012), doi:10.1111/j.1365-2966.2012.21473.x, arXiv:1204.3674.

[189] S. M. Croom, T. Shanks, B. J. Boyle, R. J. Smith, *et al.*, *The 2dF QSO Redshift Survey - II. Structure and evolution at high redshift*, **325**, 483 (2001), doi:10.1046/j.1365-8711.2001.04389.x, astro-ph/0012375.

[190] S. M. Croom, B. J. Boyle, T. Shanks, R. J. Smith, *et al.*, *The 2dF QSO Redshift Survey - XIV. Structure and evolution from the two-point correlation function*, **356**, 415 (2005), doi:10.1111/j.1365-2966.2004.08379.x, astro-ph/0409314.

[191] A. D. Myers, R. J. Brunner, R. C. Nichol, G. T. Richards, *et al.*, *Clustering Analyses of 300,000 Photometrically Classified Quasars. I. Luminosity and Redshift Evolution in Quasar Bias*, **658**, 85 (2007), doi:10.1086/511519, astro-ph/0612190.

[192] N. P. Ross, Y. Shen, M. A. Strauss, D. E. Vanden Berk, *et al.*, *Clustering of Low-redshift ($z = 2.2$) Quasars from the Sloan Digital Sky Survey*, **697**, 1634 (2009), doi:10.1088/0004-637X/697/2/1634, arXiv:0903.3230.

[193] Y. Shen, M. A. Strauss, N. P. Ross, P. B. Hall, *et al.*, *Quasar Clustering from SDSS DR5 : Dependences on Physical Properties*, **697**, 1656 (2009), doi:10.1088/0004-637X/697/2/1656, arXiv:0810.4144.

[194] J. da Ângela, T. Shanks, S. M. Croom, P. Weilbacher, *et al.*, *The 2dF-SDSS LRG and QSO survey : QSO clustering and the L-z degeneracy*, **383**, 565 (2008), doi:10.1111/j.1365-2966.2007.12552.x, astro-ph/0612401.

[195] S. M. Croom, G. T. Richards, T. Shanks, B. J. Boyle, *et al.*, *The 2dF-SDSS LRG and QSO Survey : the spectroscopic QSO catalogue*, **392**, 19 (2009), doi:10.1111/j.1365-2966.2008.14052.x, arXiv:0810.4955.

Chapitre 2

The SDSS-IV eBOSS quasar sample

Contents

This thesis work uses a 2-year data taking sample of quasars from eBOSS, one survey of the fourth generation of the Sloan Digital Sky Survey (SDSS). Section 2.1 provides a brief overview of the different generations of the SDSS and their main surveys, then section 2.2 presents the main technical characteristics of the Sloan telescope, the camera and the spectrographs that were designed for BOSS and also used for eBOSS. The astrophysics of quasars is presented in section 2.3, this section is heavily influenced by the lectures on *Trous noirs super-massifs, noyaux actifs et quasars* by Francoise Combes that I attended during the first year of my PhD at the College de France and by lectures at the first eBOSS collaboration meeting I attended on March 2016. In section 2.4, we focus on the eBOSS observation strategy which is in two steps, first a photometric survey and then a spectroscopic survey with a dedicated selection for quasar targets and a spectroscopic pipeline that classifies and assigns a redshift to each object. Section 2.5 provides further details on the redshift determination for quasars and its impact on clustering. Then, section 2.6 explains the creation of the large-scale structures catalogs used as input for clustering analysis. Eventually, in section 2.7, we present the methodology to estimate the two-point correlation function from the catalogs.

In addition to the technical articles mentioned in the text, references for this chapter include the SDSS website [1] where data releases are made public, the SDSS Project Book [2] and Ann Finkbeiner's book *A grand and bold thing* [1].

2.1 The Sloan Digital Sky Survey

The Sloan Digital Sky Survey started being named this way in 1991 when the Sloan fundation decided to support the project led by Jim Gunn, Rich Kron and Donald York to build an instrument capable of taking the spectra of hundreds of galaxies at once in order to obtain the largest map of the structures in the universe. Jim Gunn imagined the design of the required camera few years before with the conclusion that a 2.5-meter telescope would be enough for the imaging system he developed. Thanks to the latest development of coupled-charge devices, CCDs, along with the fast growing data processing capabilities, large and deep sky surveys like the SDSS have been made possible. Not only did the SDSS change the way astronomers took data, but it also changed the way astronomers worked. The SDSS marked the beginning of big collaborations in astronomy with the ultimate goal to map the universe at all scales, from stars in the Milky Way to the biggest clusters of galaxies.

2.1.1 SDSS-I and II

SDSS [2] first light was obtained on May 1998 then observations started in 2000. The first two generations of SDSS were completed from 2000 to 2005 and 2005 to 2008 respectively whose main results were :

— A multi-band photometric survey (u,g,r,i,z) of 11 663 deg^2

— A contiguous imaging and spectroscopic sample over 7500 deg^2 of the Northern Galactic Cap, called the *Legacy Survey* with more than 930 000 galaxies and 120 000

1. http://www.sdss.org
2. https://www.astro.princeton.edu/PBOOK/

quasars. This sample represents the seventh data release of SDSS which is detailed in [3].

— An imaging and spectroscopic survey over 3500 $\deg^2$ at lower Galactic latitudes to explore the structure, composition and kinematics of the Milky Way. This program is called *SEGUE* (Sloan Extension for Galactic Understanding and Exploration, [4]) and contains a sample of 460 000 stars.

— A repeat imaging of $> 250\,\deg^2$ on the Celestial Equator to discover Type 1a SN with $0.1 < z < 0.4$, called the *SDSS Supernova Survey* which discovered almost 500 spectroscopically confirmed Type 1a SNe to learn more about the cosmic acceleration [e.g. 5]

By mid-2004, there were 400 papers writen by SDSS members and 125 by non-SDSS members using Sloan data. In the fall of 2009, 2,656 papers were based on Sloan data and were cited in other papers 100,000 times. SDSS had definitely reached a milestone, with a significant impact on a broad range of subjects in astrophysics and cosmology. One of the most famous results from SDSS was the first detection of the BAO signal in the two-point correlation function of a sample of around 46,000 LRG [6] at an effective redshift of 0.35 which allowed a 5% measurement of the volume-averaged BAO distance, D_V as mentioned in the previous chapter.

2.1.2 SDSS-III

The third generation of SDSS [7] started in Autumn 2008, directly following SDSS-II. It used the same telescope but a significantly upgraded version of the double-armed spectrographs. SDSS-III was composed of four different surveys : the Baryon Oscillation Spetroscopic Survey (BOSS), the Sloan Extension for Galactic Understanding and Exploration 2 (SEGUE-2), the Apache Point Observatory Galactic Evolution Experiment 1 (APOGEE-1) and the Multi-object APO Radial Velocity Exoplanet Large-area Survey (MARVELS).

— **BOSS** was a six-year program of spectroscopic galaxy observations which has been designed to refine the measurement of the standard ruler using LRG in the redshift range $0.2 \leq z \leq 0.75$ and with enough sensitivity to constrain both the expansion rate and the angular diameter distance to percent level precision. The final data release DR12 consists of 1.2 million galaxies with redshifts between 0.15 and 0.7 with an effective area of 9329 $\deg^2$ [8]. It also includes a second program dedicated to Ly-α forest to constrain the BAO scale at $z \sim 2.5$ using the neutral hydrogen located in the intergalactic medium as a tracer of large scale structures. The final sample recorded 160,000 quasars between redshift 2.15 and 3.5 and measurements of the auto-correlation [9] and cross-correlation [10] provided constraints on the BAO distances at few percents level.
Thanks to very good weather conditions during the survey, the observations were ahead of planning and ended in February 2014. The leftover time was used for additional ancillary programs which were proposed by SDSS-III members (they cover various subjects such as high energy blazars, very low mass stars and quasar selection using their variability) and for an eBOSS pilot program known as the Sloan Extended Quasar, ELG, and LRG Survey (SEQUELS). It represented 128 plates that were used to demonstrate the data quality of each eBOSS target class and test the pipeline procedure, 66 plates were completed before the end of SDSS-III and were included in DR12.

— **SEGUE-2** produced spectra for almost 118,000 stars inside the galactic halo of our galaxy. In combination with the 230,000 spectra from SEGUE-1, the data revealed the complexity of the kinematic and chemical sub-structures of the stellar halo of the Milky Way and provided insights on the metal enrichment of our galaxy. SEGUE-2 used the telescope "dark" time between Autumn 2008 and Spring 2009, ending before the start of BOSS in Autumn 2009.

— **APOGEE-1** aimed at measuring with high precision the peculiar velocity and the chemical composition of our galaxy by observing around 100,000 red giant stars located in different regions of the Milky Way (bulb, disk, bar and halo). The observations started in Spring 2011 and ended in Spring 2014. It used high-resolution infrared spectroscopy during the "bright" time (when the Moon is more than 60% illuminated) that BOSS could not use.

— **MARVELS** was a spectroscopic survey designed to observe 11,000 bright stars of our galaxy. Each star needed to be observed between 25 and 35 times on an eighteen months period, looking at their radial velocity to look for giant gaseous exoplanets. It started in Autumn 2008 and observed during the "bright" time with the goal to constrain theoretical models of formation and evolution of giant planets systems. However, the required resolution of the spectra was never reached and the project was stopped in 2012.

2.1.3 SDSS-IV

The fourth generation of SDSS [11] started in Autumn 2014 and will end in 2020. It includes an extension of BOSS, the eBOSS program (extended-BOSS), a second generation of APOGEE, and a new survey, the Mapping Nearby Glaxies at APO (MaNGA).

— **eBOSS** is the continuity of BOSS whose goal is to obtain percent-level measurements of Baryon Acoustic Oscillations (BAO) from $0.6 < z < 3.5$, in particular in the almost unexplored redshift range $0.7 < z < 2.2$. Observations started in the fall 2014 and will finish in February 2019. It includes four classes of target : a sample of LRG at higher redshift than what of the BOSS sample, a new sample of emission line galaxies (ELG) which are blue star-forming galaxies in the redshift range $0.6 < z < 1.2$, a new sample of low-redshift quasars between 0.8 and 2.2 and additional Lyman-α forest at redshift $z > 2.2$. Details about the science forecasts can be found in [12, 13].
eBOSS also includes two sub-programs to follow up on other types of objects : the Time-Domain Spectroscopic Survey (TDSS) for variable objects (quasars and several classes of variable stars), and the SPectroscopic IDentification of ERosita Sources (SPIDERS) whose aim is to provide a complete and homogeneous optical spectroscopic follow-up of X-ray sources detected by eROSITA. The first data release of eBOSS (DR14, [14]) was made public in 2018, the analysis of the clustering of the LRG sample at an effective redshift of 0.72 is presented in [9], the quasar catalog in [15], the first BAO detection in the low-redshift quasar sample and the analysis of the full-shape two-point correlation function using quasars as tracers of the matter field at an effective redshift of 1.52 are the subject of this thesis whose results have been published in [16, 17].

— **APOGEE-2** continues to focus on observations of giant stars and uses the 2.5m du Pont telescope at Las Campanas Obervatory in Chile to observe the Southern hemisphere. It extends into previously unreachable parts of the Milky Way disk (APOGEE-1 was focused on red giant stars distributed across several kiloparsecs of the disk). APOGEE-2 also acquires spectra of young stars and star-forming regions,

variable stars, stars in clusters and satellite galaxies, and stars with asteroseismic measurements.

— **MaNGA** aims at understanding the history of present day galaxies (birth, assembly and ongoing growth) by providing two-dimensional maps of stellar velocity and velocity dispersion with measurements of stellar and ionized gas metallicity, element abundance ratio, star formation rate and dust extinction. Its goal is to measure spectra across the face of $\sim$ 10,000 nearby galaxies with no cut on size, morphology or environment in order to obtain a sample that is fully representative of the local galaxy population.

2.2 Technical characteristics

The SDSS was imagined as a multi-band photometric survey covering a large fraction of the observable sky associated with a large spectroscopic survey. As explained in [18], such requirements imply several technical innovations for each element of the experiment :

— telescope : wide field of view, very low distortion focal plane, very precise pointing, capacity to switch between photometry and spectroscopy easily

— camera : wide field of view, simultaneous observations in different photometric bands, good precision for astrometric calibration

— spectrographs : capacity to take hundreds of spectra simultaneously on a wide wavelength coverage

— acquisition system : unprecedented data storage capabilities and real time control quality

— data pipeline : fast, efficient calibration of the data and identification of objects both in photometry and spectroscopy

2.2.1 The telescope

In order to reach the goals mentioned above, a rough optimization of parameters such as telescope aperture, available field with acceptable image quality, available optical fibers, and available CCDs led to the design of a telescope of approximately 2.5 meters aperture and focal ratio (focal length of the telescope divided by its aperture) of approximately f/5, and with a field of view of 3 degrees. Such a telescope was built at the Apache Point Observatory located in New Mexico at 2800 meters above sea level. Details on the choice for the design and on the performance of the telescope can be found in [18] who designed a modified Ritchey-Chrétien telescope with a flat field by making the curvatures of the primary and secondary mirrors the same, which yields zero curvature in the focal plane. In addition to the classical Ritchey-Chrétien design, it has two optical correctors. The SDSS telescope was designed to carry out both imaging and multi-object spectroscopic observations, scheduled and interleaved depending on the quality of the weather. Not only does the telescope need to support these observing modes, but it also allows rapid change between imaging and spectroscopy.

The left panel of figure 2.1 shows a picture of the SDSS telescope at Apache Point Observatory. The right panel represents a schema of the telescope with its components :

— The primary mirror has a diameter of 2.5 meters and a focal to diameter ratio of $f/2.25$. It has a hole of 1.17 meters in its center. There is a conical light baffle above it to prevent parasitic light to reach the focal plane.

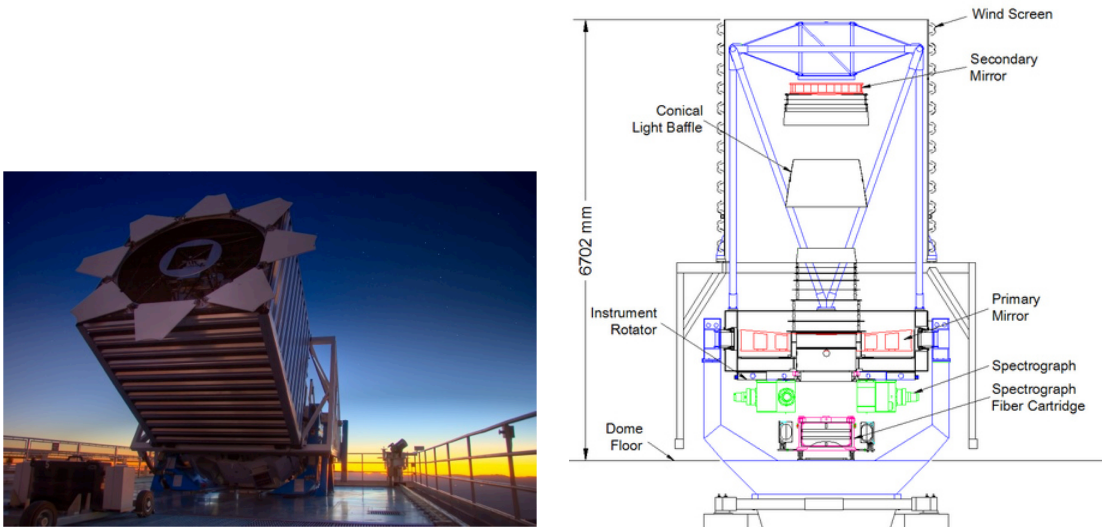


FIGURE 2.1 – Left panel : Picture of the 2.5 meters diameter telescope designed by the SDSS collaboration. Credit : http://blog.sdss.org/image-gallery/. Right panel : Schematic view of the SDSS telescope where the primary and secondary mirrors are represented in red, the spectrographs in green and a cartridge in pink. From [19]

— The secondary mirror has a diameter of 1.08 meters. It also has a light baffle, that in addtion to the primary mirror, blocks 27% of the light reaching the telescope.

— The first corrector is of Gascoigne type and it is the last optical piece common to the photometric and spectrometric configuration. That is why it is sometimes called the "common" corrector. Its goal is to drastically reduce the astigmatism which is one of the major drawbacks of Ritchey-Chrétien telescopes.

— The second corrector or "final" corrector is not the same depending on the configuration of the telescope. These configurations are optimized for each case and are briefly presented in Section 2.2.2 for the photometric corrector and in Section 2.2.3 for the BOSS spectroscopic corrector.

The two mirrors form an optical system with a focal to diameter ratio of $f/5.0$ and a 3 degrees field of view. The two mirrors are separated by only 3.6 meters, so that the focal plane is located 0.76 meter behind the primary mirror, allowing an easy access to the instruments.

2.2.2 The camera

The photometric survey of BOSS and eBOSS used the same camera that was used for SDSS and SDSS-II because it was already optimized for the operating mode and the 3 degrees field of view of the telescope. A full study of the camera can be found in [20], here we just summarizethe main elements :

— **Optical corrector** : This lens is the first piece of the camera and one of the most important as it ensures the preservation of the mechanical properties of the camera thus having a big impact on the image quality and the astrometry. It aims at correcting the distortion at the focal plane induced by the telescope optical system.

— **Matrix of 6 columns of 5 CCDs each** : There is one CCD for each photometric band : u' ($\lambda_{\text{eff}} = 3550$), g' ($\lambda_{\text{eff}} = 3550$), r' ($\lambda_{\text{eff}} = 3550$), i' ($\lambda_{\text{eff}} = 3550$) and z'

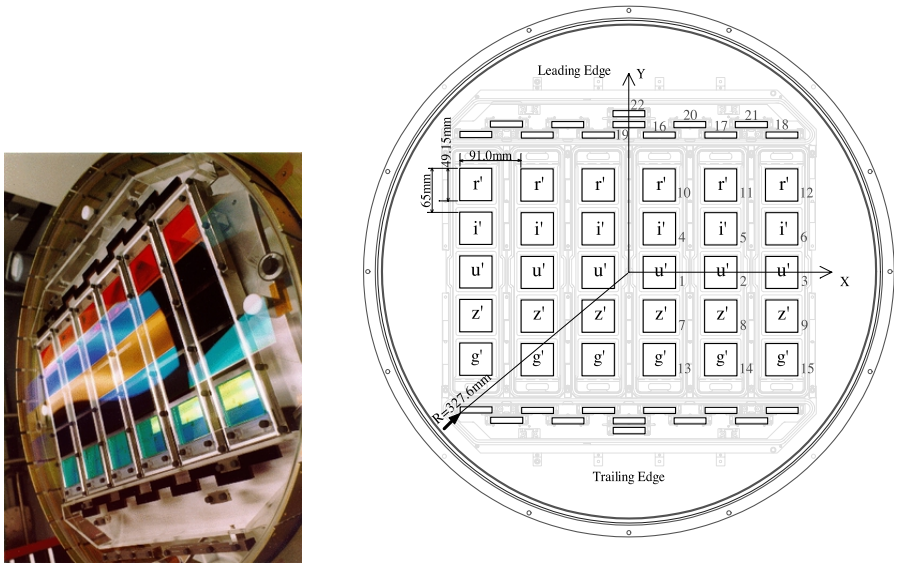


FIGURE 2.2 – Left panel : A photograph of the SDSS camera. Right panel : The focal plane organization for the camera used for SDSS, SDSS-II, BOSS and eBOSS. The matrix of photometric CCDs is at the centre, whereas the astrometric and focus CCDs are at the top and bottom. The drift scanning direction is from top (leading edge) to bottom (trailing edge). From [20]

($\lambda_{\text{eff}} = 3550$) and its size is 2048×2048 pixels. The remaining space on the focal plane is used for 24 smaller CCDs (2048×400 pixels), 22 being used for the astrometry and 2 for the focus. Indeed, the CCDs for photometry cannot measure accurately the position of bright sources because their sensors saturate. The repartition of these smaller CCDs around the bigger ones is visible in the left panel of figure 2.2. A technique called time delay integration or drift scanning is performed from top to bottom on the schema in the right panel of figure 2.2 where the telescope stands pointing in a given direction while the sky passes under the effect of the rotation of the Earth. Therefore, a star entering the focal plane successively encounters an astrometric CCD, the photometric CCDs in the order r', i', u', z', g' and eventually a second astrometry dedicated CCD. All the CCDs dedicated to the photometry are identical but they have filters to select the wavelength band of observation. The CCDs can observe from the UV atmospheric cut off around 3000 Angstrom to the limit of silicon based detectors close to 11000 Angstrom. The wavelength range and quantum efficiency (defined as the ratio between the numbers of collected electrons to the number of incident photons) of each filter is shown in figure 2.3. The very low efficiency of the u' band is the reason why it is located at the centre of the focal plane. The magnitude detection limits are defined for a signal-to-noise ratio of 5. They are approximately $u'_{\text{limit}} = 22.1\,\text{mag}$, $g'_{\text{limit}} = 23.2\,\text{mag}$, $r'_{\text{limit}} = 23.1\,\text{mag}$, $i'_{\text{limit}} = 22.5\,\text{mag}$ and $z'_{\text{limit}} = 20.8\,\text{mag}$.

2.2.3 Spectrographs

BOSS has two identical spectrographs that have been rebuilt from the original SDSS spectrographs to take into account upgrades like the increase in the number of fibers. They are thus called the BOSS spectrographs and they remain the active optical spectrographs

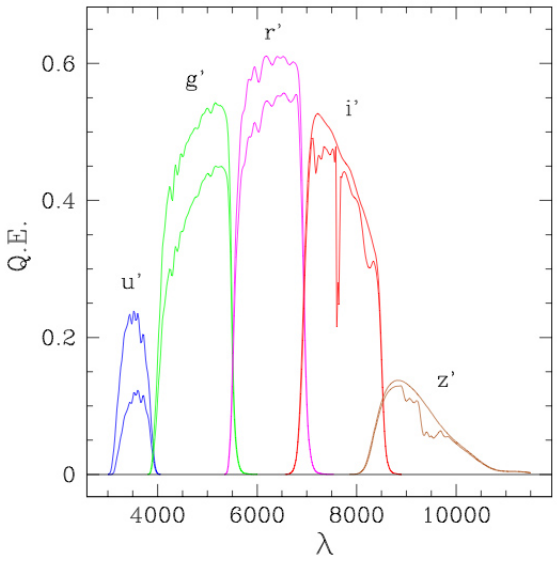


Figure 2.3 – Quantum efficiency of the optical system in each band. The upper curves take into account the transmission of the filters, the quantum efficiency of the CCDs and the losses in the optical system. The lower curves add the atmospheric extinction. From [20].

in SDSS and it was used to collect all MaNGA and eBOSS data released in DR14.

In what follows, we summarize the most important characterictics of the elements of the spectrographs for which a full study has been conducted in [19] :

— **Cartridges** : To get the spectra, an aluminium plate is put at the focal plane of the telescope. This plate is 3.2 mm thick, has a diameter of 81.3 cm and weighs 4.3 kg. The plate is drilled in advance, where optical rays from an object are going to converge, at the position of targets known from the target selection (see the next section). Each plate is observed for typically one hour, so for one night of observation that is at most 9 hours long, up to 9 cartridges are prepared to optimize the data taking. Out of the 1,000 fibers for one plate, there are 80 fibers to measure the mean flux from the sky to substract the sky background and 20 fibers are used for calibration purposes pointing at specific locations. 895 fibers are available for new targets and 5 are assigned to objects observed with at least another plate to control the reproductibility of the observations. In addition to the 1,000 holes mentioned, 16 holes are drilled to follow reference sources for pointing and monitoring the observations. Holes are also drilled at the position of bright stars to avoid parasitic reflections on the plate. Optical fibers are then plugged in these holes to redirect the light to the spectrograph grism, which is a combination of a prism and a grating. The plates and the optical fibers are held by an aluminium structure, the whole forming a cartridge that can be seen in figure 2.4. A cartridge can be easily mounted on the telescope by one person and several cartridges (9 are available) can be fiber-plugged in advance, making this system easy to operate and flexible.

— **Optical fibers** : In order to increase the statistics of the survey, the number of fibers per plate went from 640 to 1000 for SDSS-III/BOSS and eBOSS. Doing so without changing the telescope optics required to decrease the size of the fibers from 180μm (3$''$ on the sky) to 120μm (2$''$ on the sky). Each fiber is glued into a ferrule to protect

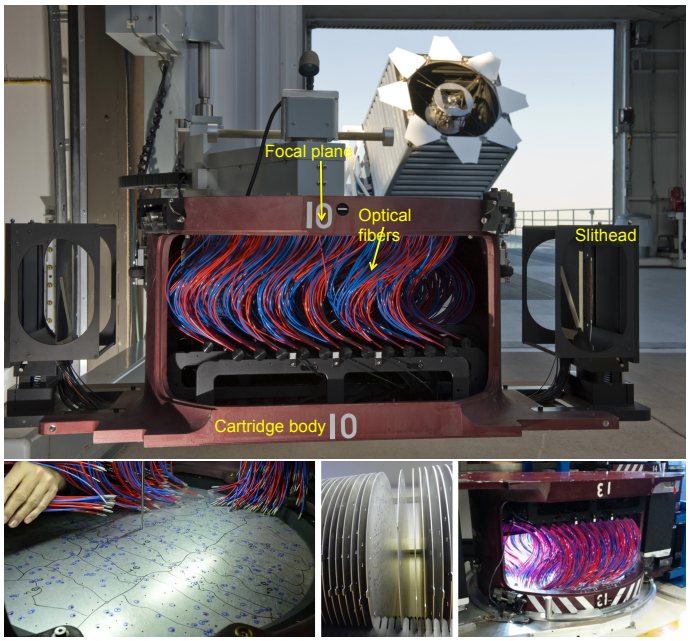


FIGURE 2.4 – Top panel : The BOSS/eBOSS fibers mounted on the cartridge in the fore-front with the 2.5-meter Sloan telescope in the background. Bottom panel : The aluminium plate being plugged with optical fibers, a set of drilled plates being ready to be installed when the cartridge (right) is mounted on the telescope. Credit : SDSS

it so that two fibers on a plate are separated by at least $62''$ on the sky. Each fiber collects the light at the focal plane in a cone of 0.1 numerical aperture [3] and gives back the light in a slightly wider cone of 0.125 n.a. This deterioration is due to the propagation of light inside the optical fiber. At the end of the fiber, the collimator is built to collect light from a cone of 0.125 numerical aperture, any ray outside this cone is lost. Thus there is a requirement to control the beam width at the output of the fibre to maximize the yield of the system. To do so, the spectrographs are mounted directly on the telescope, limiting both the displacement between the two ends of the fibres and the mechanical constraints. It also allows a minimization of the fibre length, each of them measuring 1830 ± 25mm. The 1000 fibres are divided in two sets of 500, grouped in bundles that are directed to two slits (slitheads), one at each end of the cartridge. Each slit goes into a spectrograph when the cartridge is mounted on the telescope.

— **Blue and red cameras** : Each spectrograph has two cameras : a blue one from 3600 to 6350 Angstroms (the blue limit is set to include the wavelength-calibration Cd I arc line at 3610.51 Angstroms) with a resolution power between 1560 and 2270, the red camera covers $5650 < \lambda < 10\,000$ Angstroms (the red limit is set by the use of silicon detectors) with a resolution between 1850 and 2650. The resolution power R is defined by $R = \lambda/\Delta\lambda$ where $\Delta\lambda$ is the FWHM of the point spread function [4] (PSF) of the CCD image by the spectrograph. It depends on the wavelength and on the position in the focal plane of the CCD, i.e. to the fiber number. As shown in figure 2.5, for all CCDs the resolution is lower on the sides, at the edges of the

3. The numerical aperture of a medium is defined as $n.a. = n \sin\theta$ where n is the index of refraction of the medium that light travels and θ is the maximum possible angle of a light ray that enters the medium.
4. The PFS is the response of an imaging system to a point source, i.e. a Dirac function.

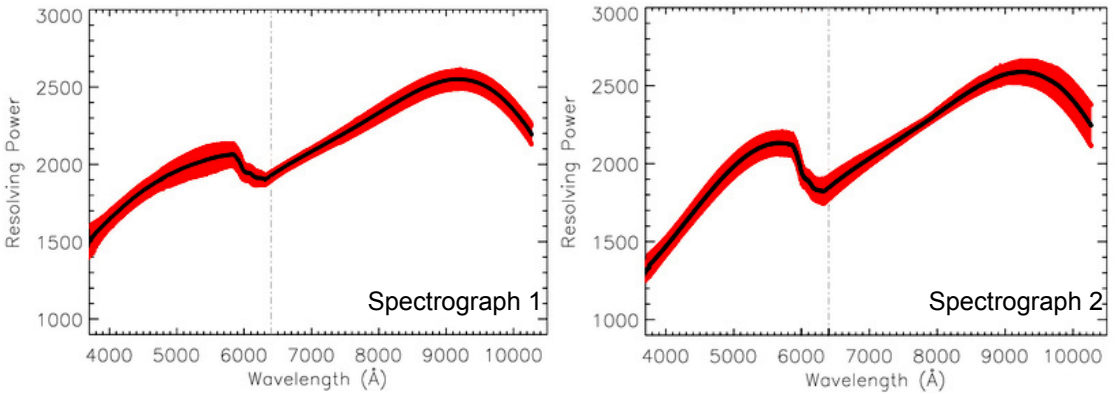


Figure 2.5 – The resolution power for the BOSS spectrographs. The red regions correspond to the regions that contain 68% of the light. The dashed line shows the transition between the blue CCD and the red one. From [19].

spectrographs, by a factor of 2 compared to the center. We will see that this behavior has an impact in the redshift efficiency, i.e. the ability to provide a valid redshift when a spectrum is acquired.

A scheme of the optical system of a spectrograph is presented in figure 2.6. Light arrives in the spectrograph through a slit located at the end of a cartridge, where the slit holds 500 fibers. First, the light goes on a collimator which reflects it as a parallel beam to a beamsplitter oriented at 45 degrees. At the beamsplitter, wavelengths shorter than 6400 Angstroms are reflected in direction of the blue channel whereas longer wavelengths are transmitted to the red channel. In each channel, the beam is diffracted by the grism which decomposes the light of each fiber. Each component is observed by cameras located after the grisms and which are made of a succession of 8 lenses.

2.3 eBOSS quasars

eBOSS is mapping the structures of the universe using four cosmological tracers that complement the sample of BOSS galaxies and Ly-α forests. Figure 2.7 shows these tracers and the redshift range they allow us to probe. In this thesis, we use the quasars themselves as direct tracers of the matter density field to perform a clustering analysis. By definition, a clustering analysis is a statistical analysis of the distribution of pairs of galaxies so that it involves scales above the typical scale of galaxies (above few Mpc). In general, it is not necessary to understand the astrophysical processes that are associated. Nevertheless, quasars are fascinating objects which are not very-well known, so it is worth dedicating the next section to present them.

Thanks to the development of radio-astronomy in the 1950-1960's, dedicated telescopes have been built to receive radio waves from the sky. A growing number of optical point-source objects with radio emission have been discovered, and the term quasar has emerged in May 1964 by the astrophysicist Hong-Yee Chiu in *Physics Today*, to describe these puzzling objects : "So far, the clumsily long name *quasi-stellar radio sources* is used to describe these objects. Because the nature of these objects is entirely unknown, it is hard to prepare a short, appropriate nomenclature for them so that their essential properties are obvious from their name. For convenience, the abbreviated form 'quasar' will be used throughout this paper."

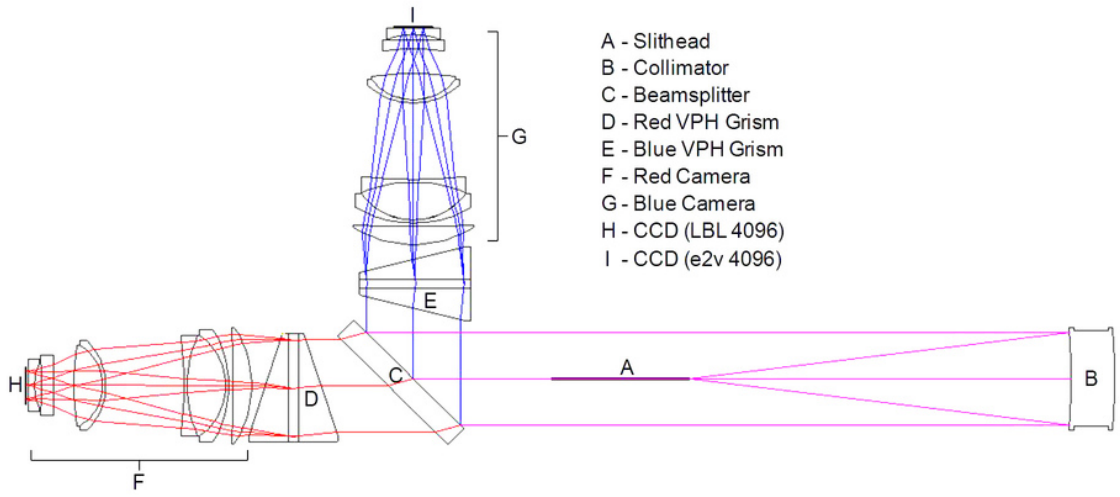


FIGURE 2.6 – Optical schema of the BOSS spectrograph : light arrives in the spectrograph through the slithead (A), then it is directed to the collimator (B) which returns it in the opposite direction as a parallel beam. The beamsplitter (C) splits the beam into a red and a blue components, each component is diffracted by a grism (D and E) and then focused by the camera (F and G). From [19].

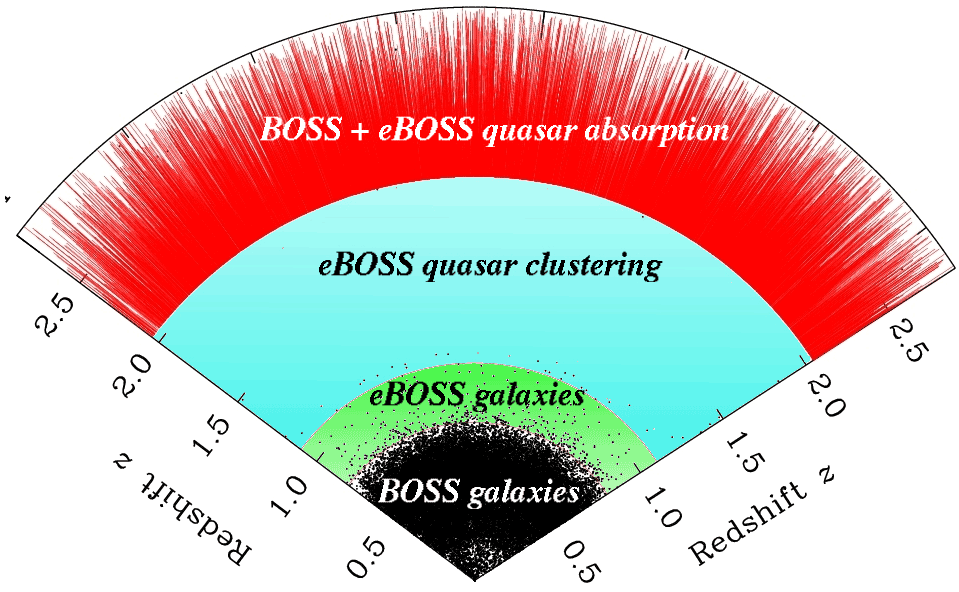


FIGURE 2.7 – BOSS and eBOSS cosmological probes as a function of redshift. This work uses the eBOSS quasar sample between $0.8 \leq z \leq 2.2$. Credit : SDSS eBOSS

2.3.1 Surprising radio sources

If it is true that the first quasars were confirmed in 1963, actually radio sources have started being discovered since 1936 by Grote Rober, an amateur astronomer who built a radio telescope in his backyard. Over the next several years, he found many radio sources, among them Cygnus A which is now one of the strongest known radio sources. Fifteen years later, in 1951, astronomers Walter Baade and Rudolph Minkowski found the object that created Cygnus A's radio emissions : they used the 200-inch visible-light telescope on Mount Palomar in California to find an unusual-looking faint optical galaxy [21]. But when they looked at the spectrum of the galaxy, they found an even greater surprise. Cygnus A turned out to be a galaxy with a redshift of 0.057. This redshift measurement put it over 700 million light years from us, which made it the most distant observed object and the most intense radio source at that time. Then, hundreds of radio sources with no corresponding visible objects, very small angular size and some optical counterparts were recorded in the Third Cambridge Catalogue which was published in 1959 and produced using the Cambridge Interferometer [22]. In 1963, two objects in this catalog have been confirmed to be associated with optical objects : the radio source 3C48 by Allan Sandage and Thomas A. Matthews [23] and confirmed in by Jesse L. Greenstein and Maarten Schmidt together with the discovery of the radio source 3C273 [24]. Thanks to a series of measurements taken by Cyril Hazard and John Bolton during occultations using the Parkes Radio Telescope, Maarten Schmidt could identify the object and obtain its optical spectrum using the 200-inch Hale Telescope on Mount Palomar. This spectrum, so as the one of 3C48, revealed strange emission lines which were actually spectral lines of hydrogen that were redshifted. Schmidt and Greenstein concluded that the radio source 3C273 was identified with an object at a redshift $z = 0.158$. The optical luminosity implied by this redshift is about 120 times the magnitude of an object like the Andromeda galaxy.

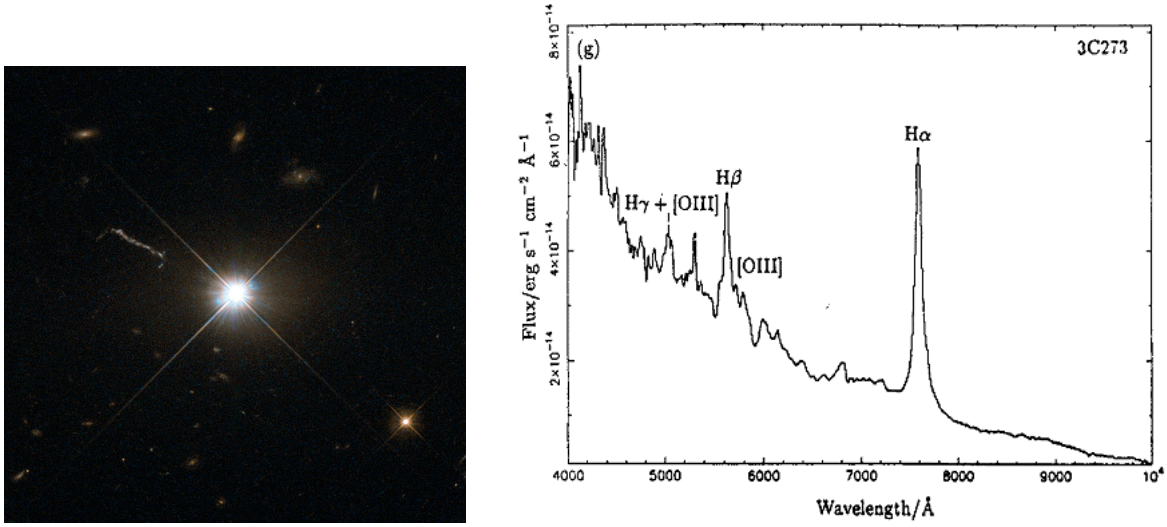


Figure 2.8 – Left panel : Optical image of the quasar 3C273, credit : HST. Right panel : Optical spectrum of the quasar 3C273.

In optical photographic plates, the object looked like a star, as one can see on the left panel of figure 2.8 which corresponds to the optical image of 3C273 from the Hubble Space Telescope [5], while its optical spectrum is shown on the right panel. Quasars typically have

5. http://hubblesite.org

optical-ultraviolet spectra and broad emission lines, sometimes with widths suggesting Doppler motions of thousands of km.s^{-1}. The optical-ultraviolet continuum varies on timescales of weeks or less, suggesting physical sizes of light-weeks.
Since the last twenty years, we have discovered about half a million of radio source objects, not all of them are quasars as we will see in the following section that there is an important diversity among these objects.

2.3.2 Quasars and active galaxies

When Greenstein and Schmidt published their study of 3C48 and 3C273 in 1996, they wrote at the beginning of their abstract : "Together with the radio-frequency data and the light variability, these indicate the presence of very large total energies in a relatively small volume of space. We deliberately have not attempted to discuss the origin of these large energies, nor do we discuss the numerous other physical problems concerned with suggested mechanisms in the quasi-stellar objects."

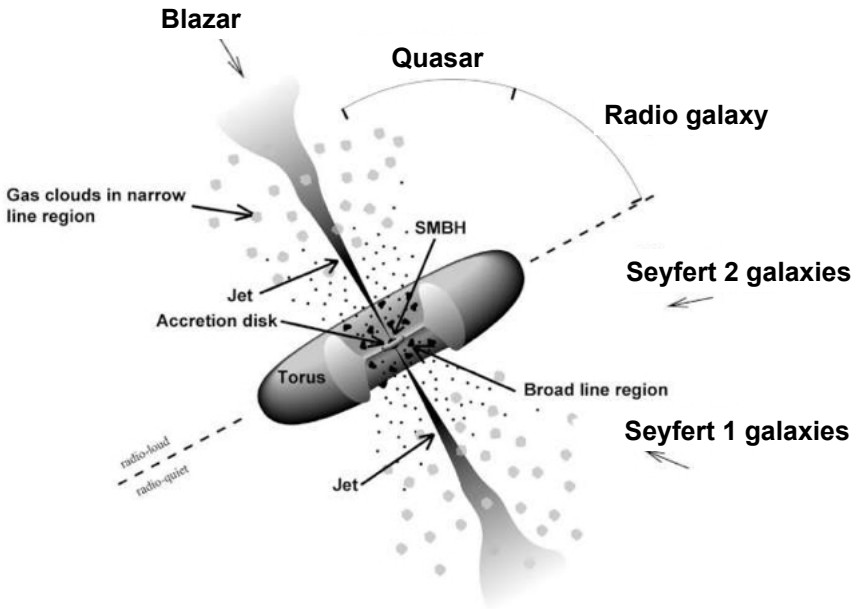


FIGURE 2.9 – Schematic view (not to scale) of the dust-torus-based unified model of AGN where the orientation according to which we see the active galaxy determines its class. Credit : Fermi collaboration.

In fact, quasars belong to a more general class called active galactic nuclei (AGN). The mystery of AGN is that they produce very high luminosities in a very concentrated volume, which requires a huge amount of gravitational energy. In the 1960-1970's, a theory emerged to explain this extraordinary amount of gravitational energy that is associated with quasars [e.g., 25, 26] : they are powered by super-massive black holes (SMBH) of billions of solar masses as illustrated in figure 2.9. The radiated energy we receive comes from matter being accreted onto the black hole. The matter surrounding the central black hole is likely to form an accretion disk, in which some source of viscosity drains the orbiting matter from angular momentum, making it spiral inwards toward the central black hole. AGN can be classified depending on their luminosity and radio-loudness :

— **Luminosity** : Quasars are the most luminous AGN and Seyfert galaxies are generally less luminous than quasars with a luminosity a hundred times the one of the Milky

Way, while quasars have generally luminosities between $10 - 1000$ times the one of the Milky Way. The former have been discovered in 1943 by Carl Seyfert who found that some spectra presented very broad emission lines and others much narrower lines just like quasars did [27].

— **Radio-loudness** : Sometimes, ejections of energetic particles, called outflows, occur along the poles of the disk, escaping and forming collimated radio-emitting jets. Quasars can be divided into two groups : radio-loud quasars (also called QSO) which have radio-emitting lobes and which correspond to only $\sim 10\%$ of all the quasars and radio-quiet quasars. In what follows, quasar refers to both. It is still not clear why some galaxies have radio lobes while others do not, but the lobes appear to be caused by particle jets that are evicted from the central engine. These outflows interact with the interstellar and intergalactic mediums, creating a plasma in which electrons spiral along magnetic fields lines and emit synchrotron radiation at radio wavelengths.

All the different types of AGN have been unified in a common picture where the optical appearance depend strongly on the orientation as shown on the top panel of figure 2.9. When the central black hole accretion (often called the central engine) is visible along the LOS, the broad emission lines and blue optical-ultraviolet continuum are visible. When the host galaxy is visible, the corresponding AGN are dubbed Seyfert 1 galaxies ; if the galaxy is not visible they are dubbed radio-loud quasars. When the dusty torus obscures the LOS to the central engine, only narrow emission line are visible and the continuum is dominated by starlight from the host galaxy, these are Seyfert 2 galaxies or radio-quiet quasars. AGN with jets that are coincidentally pointed at the Earth are extremely bright at all wavelengths, from radio to gamma rays, due to Doppler beaming. These types of AGN are known as blazars and lots of them are observed using the Fermi Gamma Ray Space Telescope [6] and the High Energy Stereoscopic System (H.E.S.S.) [7]. Therefore, in this 'orientation-based unified model', the differences between the AGN arise simply because of the different orientations seen by the observer.

2.3.3 Quasar spectrum

Quasars emit radiations from γ-rays and x-rays to the far infrared ($\lambda \sim 100\mu$m). For most of them, the amount of energy emitted in each band is remarkably similar, in contrast to thermal radiation from stars, which is much more peaked and restricted in wavelength. Most quasars at $z \leq 2.5$ are bright at ultraviolet wavelengths, a property that is very helpful to distinguish them in sky surveys from the more numerous stars which are usually faint at these wavelengths. The continuum emission in quasars appears to arise from a combination of thermal and non-thermal processes. In any event, the continuum radiation from quasars demonstrates that some very energetic processes are involved. Furthermore, the continuum radiation at the highest energies tends to show the most variability and the shortest timescales, which is another indication of the extreme conditions that exist near quasars. The strongest emission lines in quasar spectra come from hydrogen, carbon and magnesium, with lines of nitrogen, oxygen, iron and other elements also being visible. The observed levels of ionization range from neutral for hydrogen and oxygen to five times ionized oxygen and even more highly ionized iron. A quasar spectrum is shown in figure 2.10 that we can describe in more details in the framework of SMBH :

— Radio : synchrotron emission of electrons that spiral along magnetic fields lines

6. https://fermi.gsfc.nasa.gov
7. https://www.mpi-hd.mpg.de/hfm/HESS/

— Infrared : thermal emission from the dusty torus

— Optical-UV : thermal emission of the accretion disk surrounding the central SMBH (continuum) + broad emission lines (between $\sim 3000-10,000\,\mathrm{km.s^{-1}}$) are associated with the rotating gas located at ~ 1 pc away from the black hole in the Broad Line Region (BLR) + narrow emission lines (between $\sim 200-900\,\mathrm{km.s^{-1}}$) to a region of ~ 100 pc radius around the black hole, in the Narrow Line Region (NLR).

— X–ray : low energy photons scattering with relativistic electrons that are located in a hot corona above the black hole

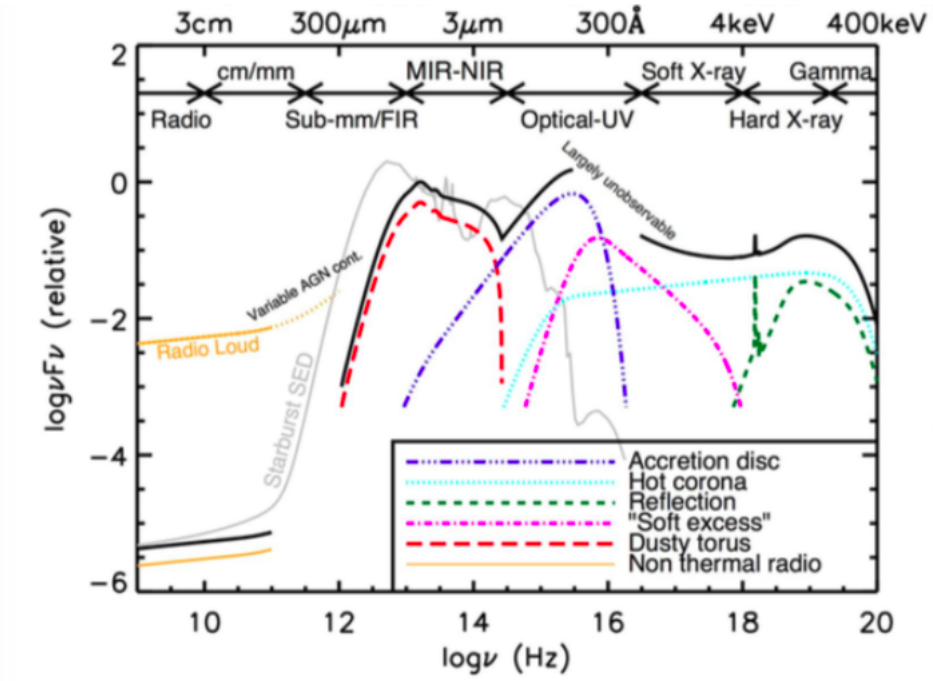


FIGURE 2.10 – Representation of a quasar spectrum in all wavelenghts.

At the end of the 1990's, correlations between the masses of the AGNs and galaxy properties were highlighted using a sample of galaxies with HST photometry and ground-based kinematics : the formation of SMBH and their growth through accretion are closely related to the formation of stars in their host galaxies [e.g., 28, 29, 30]. AGN have become a key element in the galaxy evolution paradigm, especially because they are present at the center of all galaxies [31]. Not all AGN in the center of galaxies are active quasars, however people realized that quasars could be just a phase in the normal cycle of galaxy and could therefore be used as cosmological tracers, just like galaxies. The clustering of quasars as a function of redshift can also provide useful constraints on our understanding of galaxy evolution : the large-scale clustering amplitude increases with the mass of the dark matter halos hosting the quasars. Moreover, comparison of the abundance of such halos to that of quasars can provide constraints on the duty cycle (the fraction of active quasars that reside in dark mater halos) and degree of scatter in the observable halo-mass relation. Such studies have been led using the quasar sample of SDSS-III BOSS [32, 33] and a preliminary sample of SDSS-IV eBOSS quasars [34], they confirmed the picture that quasars in the eBOSS redshift range reside in dark matter halos of mass $\mathrm{M} \sim 10^{12.5}\mathrm{M}_{\odot}$. Precise bias and mass measurements for quasars at multiple redshifts are essential to understand whether their formation could be the result of major merging of gas-rich galaxies [35, 36, 37] and

more generally to help understanding the relationship between quasars, the host galaxies and dark matter halos [e.g. 38]

2.4 eBOSS survey strategy

eBOSS has been designed to obtain a ∼3-4% measurement of the BAO distance using quasars at low redshifts ($0.9 < z < 2.2$) with an effective redshift $z_{\rm eff} = 1.52$, ∼2% with new LRG at an effective redshift $z_{\rm eff} = 0.72$ in combination with the $z > 0.6$ sample of BOSS galaxies, ∼3-5% with ELG at an effective redshift $z_{\rm eff} = 0.87$ and an improvement of a factor ∼1.4 relative to BOSS using new Lyman-α forests at $z > 2.1$. Details regarding the techniques used to forecast requirements can be found in [12, 13]. These science forecasts set the survey requirements in terms of area to be covered and density of targets to be observed. The survey strategy can be divided into two main steps : first the photometric survey where eBOSS used the imaging data that were acquired during the previous generations, then the spectroscopic survey where a dedicated technique has been developed to select quasars in the redshift range $0.9 \leq z \leq 2.2$ (this sample is called the *CORE quasars*) in order to reach the desired target density and eventually each target needs to be assigned to a plate in a process called the tiling. The specificity of each target thus affects the target selection only, the other steps are common.

2.4.1 Photometric survey

All eBOSS quasar targets come from SDSS-I/II/III images collected in the *ugriz* system [39] using the wide-field imager [20]. Quasar targets are selected for eBOSS over the same areas as BOSS and targeting for eBOSS is conducted using the SDSS imaging that is calibrated to be tied to the Pan-STARRS photometry [40] and the Wide-Field Infrared Survey Explorer [41, WISE,] bands W1 and W2 in the mid-infrared which are centered on 3.4μm and 4.6μm.

The major factors that need to be addressed in specifying parameters for the photometric survey are angular resolution (pixel size), field size, exposure time, and exposure strategy. All the photometric data were taken under *photometric conditions*(no clouds and an atmospheric extinction depending only on the airmass), with the Moon under the horizon and a seeing better than $2''$ in the r' band. Using the drift scanning technique, the camera integrates the light of an object from the moment it enters the field of view to the moment it leaves it, the CCDs being read out at the sidereal rate [8] as the sky drifts by. This enables to use more than 90% of the available observing time for actual observations. The exposure time for each object is 55s in each band. Because of the very large field of view of the telescope (3 degrees) it is necessary to operate the drift scanning along great circles to avoid transit-time differences across the imaging CCD array.
The images are then processed through a series of pipelines that determine an astrometric calibration [42], detect and measure the brightness, positions and shapes of objects [43, 44]. The photometric calibration were done by tying to photometric standard stars and by using the overlap between adjacent imaging runs to tie the photometry of all the imaging observations together ('ubercalibration', [45]). eBOSS takes advantage of the application of the ubercalibration technique to Pan-STARRS by [46] and residual systematic errors in the calibration are at the sub-percent level on all photometric bands [40].

8. 15.04 arcsecs/second

2.4.2 Spectroscopic survey

Because taking a spectrum requires a substantially longer exposure time, it is not yet possible to have a spectrum for every object that has been identified in the photometric survey. Thus the first step of the spectroscopic survey is to select the targets for which one wants the spectra. The second step is the tiling, which aims at optimizing the observations of the selected targets in order to minimize the total observation time and thus the overall duration of the project. Then, it is possible to start the observations.
In order to cover the expected 10,000 deg^2, the eBOSS survey produces 1000 spectra per 15 minute exposure and the spectroscopic observations are repeated until reaching a signal-to-noise ratio $(S/N)^2 \geq 22$ at i-band magnitude for the red camera and $(S/N)^2 \geq 10$ at g-band magnitude for the blue camera. Each 15 minute exposure is processed immediately on-site using a fast, automated data reduction pipeline to check quality.

2.4.2.1 Quasar target selection

To reach the science forecasts [12, 13], it implies several requirements for quasar target selection given the instrument capabilities. In particular, to obtain about 500, 000 quasars over a surface area of ∼7500 deg^{-2}, it requires a quasar density >58 deg^{-2} where the total density of assigned fibers is of <90 deg^{-2}. The effective target density is of ∼ 115 deg^{-2} but taking into account the targets with existing good spectroscopy from earlier generations of SDSS that are not re-observed saves 25 fibers per deg^2. The eBOSS quasar target selection is fully described in [47].

The photometric survey gives access to the photometric properties of each object, i.e. their apparent magnitude in each band *ugriz*. This apparent magnitude m is related to the flux emitted by the object f compared to the flux emitted by a reference f_0 such that :

$$m = -2.5 \log_{10} \left(\frac{f}{f_0} \right) \tag{2.1}$$

where different systems of magnitude exist, the SDSS system is close to the most common, the AB system when $f_0 = 3631$ Jy [9]. We also use the *color* of an object which corresponds to the logarithmic of the ratio of fluxes between two bands, which is equivalent ot the difference between magnitudes of the same object in two bands.

Since the discovery that most of the unresolved extragalactic objects that are bluer than the stellar main sequence are quasars [48], the UV-excess has been used as a criterion to select quasars by applying optical color cuts and selecting objects using their positions in different color-color planes. The color cuts are defined using a training sample of known objects. Figure 2.11 represents the color-color diagram, here the (u-g) vs (g-r) plane, with the positions of stars for $18.0 < g < 19$ (left panel) and $21.0 < g < 22.0$ (right). Stars tend to align in a more or less straight feature, called the *stellar locus*, and we observe the same feature for quasars where the *quasar locus* is represented by the colored line and shows the positions of quasars in the (u-g) vs (g-r) plane as a function of redshift. We can see that the overlap between both loci for quasars at $z > 2.7$ which makes their selection harder when just applying color cuts. The color-cuts technique can target quasars mainly in the redshift range 0.5-2.5 and has been used by several surveys, such as the Large Bright Quasar Survey [49], the 2dF QSO Redshift Survey [50] and the 2dF-SDSS LRG and QSO survey [51]. The quasar target selection of SDSS-I and II extended the optical

9. Jansky, 1 Jy = 10^{-26} W.Hz^{-1}.m^{-2}

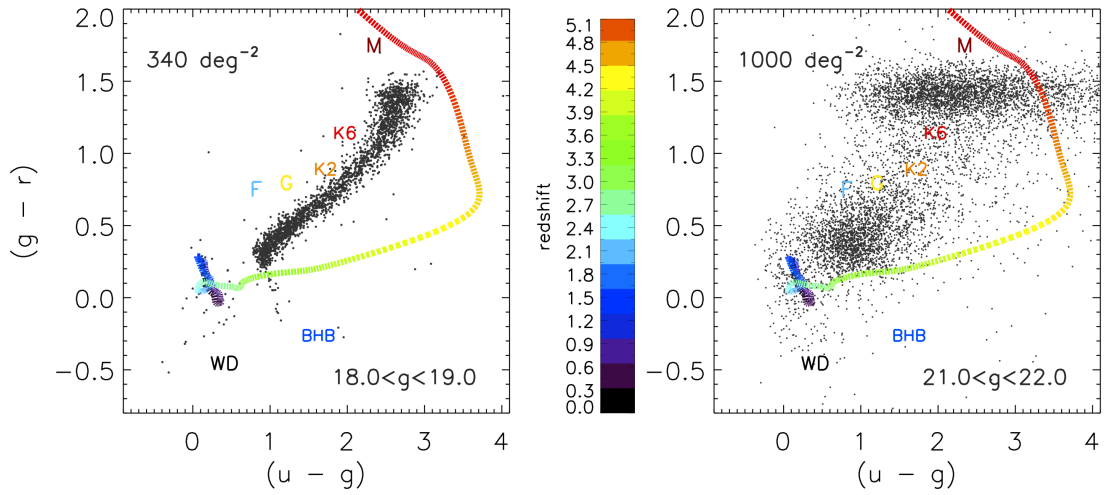


FIGURE 2.11 – Positions of stars and galaxies in the (u-g- vs (g-r) plane for $18.0 < g < 19$ (left panel) and $21.0 < g < 22.0$ (right). The positions of quasars along the quasar locus (colored line) is shown as a function of redshift. For $z > 2.7$, the stellar and quasar locii overlap, as a consequence, it is harder to select high-redshit quasars using color cuts.

cuts to quasars in the color space beyond the stellar locus up to $i < 19.1$, close to the detection limit of SDSS where the uncertainties in the measurement of fluxes are higher. Moreover, as we said, targeting quasars at $z \sim 2.7$ is challenging with color cuts because of the overlap between the quasar and stellar loci.

For all these reasons, another selection technique has been used for BOSS based on the XDQSOz algorithm of [52] in order to detect quasars at $2.1 \leq z \leq 4$ to a magnitude limit of $g < 22$ or $r < 21.85$. Details on the BOSS quasar target selection can be found in [53]. The XD (*Extreme Deconvolution*) algorithm is a method of classifying quasars in flux-space by modeling the density distribution of quasars as compared to non-quasars. Using a training sample of known objects, the algorithm is able to estimate the probability that a target belongs to a defined class, for each class : stars, galaxies, quasars at low redshift ($z < 2.2$), quasars at high redshift. Uncertainties in flux measurements are directly taken into account in the algorithm when assigning a probability. The technique also enables to distinguish the effect of missing data compared to noisy data on quasar probabilities. Figure 2.12 shows the contours at 1,2 and 3σ of the XDQSOz density probability for stars and quasars. The algorithm reproduces well the locus of each class.

For BOSS, additional methods have been developed to select quasars from photometric surveys, including the use of quasars variability [e.g. 54] and the use of multiwavelength matches (with X-rays and radio, for instance). For eBOSS quasar targeting, first we selected all point-source objects -which removed most of the galaxies- up to $g < 22$ or $r < 22$, then the XDQSOz method was extended to provide probabilistic classifications for quasars in any specified range of redshift. The eBOSS quasar selection also takes advantage of a mid-IR-optical color cut using the *WISE* imaging. The combination of both selections was tested on the pilot survey, SEQUELS and showed that the criterion that the probability to select a quasar is above 20%, i.e. $P_{XDQSOz}(z > 0.9) > 0.2$, fulfilled the requirements for the eBOSS CORE sample of > 58 deg^{-2} for quasars between $0.9 < z < 2.2$. For the clustering analysis, we extended the lower bound of the redshift range to 0.8 as we can securely select those redshifts but at lower redshifts morphological cuts affect the sample selection and the BAO signal is better sampled by galaxies. At redshift $z > 2.2$,

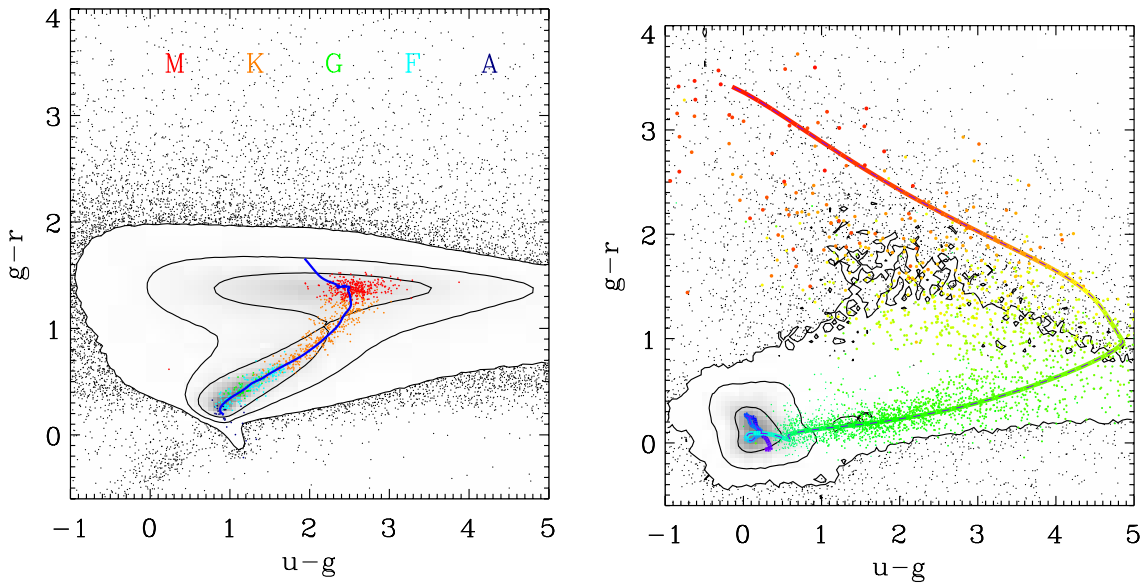


FIGURE 2.12 – Contours at 1,2 and 3 σ of the XDQSOz density probability for stars (left panel) and quasars (right panel).

the redshift measurement is less secure because some characteristic emission lines such as CIV and MgII are too redshifted to be in the wavelenght coverage of the BOSS cameras. Moreover, the high-redshift quasars are used for Ly-α forests'measurements, so we decided to cleanly separate the two volumes used for BAO measurements. Thus, our study uses spectroscopically-confirmed quasars with $0.8 \leq z \leq 2.2$.

2.4.2.2 Tiling

Once all targets (quasars, but also galaxies and targets from ancillary programs) have been determined, their angular coordinates are put into a catalog, and each target needs to be assigned to a fiber in a plate. This operation is called *tiling* and the procedure is described in [55]. The goal is to minimize the number of required plates while maximizing the number of objects that will be assigned to a fiber, providing that a plate contains 1,000 fibers, plates are circular so when a target is located in a region that overlaps two plates, one needs to decide which plate the target belongs to. An additional difficulty comes from the inhomogeneity of the target density, meaning that the centers of plates need to be closer to each other where the density is higher.
Do to so, the spectroscopic observations are split into distinct areas of the sky called *chunks*. A chunk consists of a set of rectangles in a spherical coordinate system and it is the basic unit of sky input to the tiling algorithm. Then, the tiling provides a cenrtal location for each target and a list of targets per plate. Both constitute the spectroscopic tile which is the ouput of the tiling algorithm.

2.4.3 Data reduction

The data reduction software works plate by plate where the raw data are sent from the site of the observatory to a Science Archive Server (SAS) system at the University of Utah each morning. When the transfer is complete, they are fully processed by an automated data reduction pipeline that takes roughly 10 hours to reduce a full night of data.

2.4.3.1 Pipeline

The data reduction pipeline is fully described in [56]. It consists of two largely independent steps : extraction and classification. To do so, the BOSS pipeline relies on templates that span both the full space of physical object types within the survey and the full wavelength coverage of the spectrographs. Each template is determined using a training sample of objects with known redshift and whose spectra are decomposed into a common basis of principal components (Principal Component Analysis, PCA). The quasar redshift templates were generated from a training sample of targets selected by the SDSS DR5 quasar catalog [57] and the first four principal components are used as the linear basis set for the automated redshift and classification measurements.

(i) **Extraction** : The extraction step from the raw CCD images results in wavelength-calibrated and flux-calibrated spectra. It deals with the individual exposure and transforms the two dimensional images of the CCDs in one dimensional spectra. Those spectra are then calibrated using the arc lamps data of the plate. The per pixel variance is estimated from the readout noise and the number of photons recorded in each pixel. The inverse variance is then multiplied by a factor that takes the different known flaws of the CCDs into account. The pixels that were hit by a cosmic ray are identified and masked. Finally, the flux is calibrated using the spectra of standard stars that were observed on the plate for this purpose.

(ii) **Classification** In the second step, the one-dimensional spectra are classified into object types and redshift and recorded in a catalog. It consists in adding the different individual exposures to obtain a coadded spectrum for each fibre. Each spectrum has the data of both the blue and red CCDs, covering the full range of wavelength of the instruments $3650 - 10400$Angstrom. The spectra are then rescaled to be linear in $\log(\lambda)$. The variance is estimated using the variance of each exposure, the covariance being neglected. For objects that have been observed multiple times, the different spectra are compared and the spectrum with the best signal to noise ratio is kept as the primary spectrum. When the quality of a spectrum is too bad, it can be analyzed by the pipeline because it will yield an unreliable classification and unsecure redshift. A 'ZWARNING>0' flag is assigned to those spectra.

The top panel of figure 2.13 shows two eBOSS quasar spectra that have been taken under good observational conditions ($g \sim 20$) while the two quasar spectra on the bottom have been observed at the limit of the g-band magnitude detection ($g \sim 22$) so the spectra are noisier and the identification of emission lines is more difficult. The red curve corresponds to the pipeline model and the green curve is the sky background whose contribution is removed by the pipeline.

A last step consists in assigning a redshift to each object. This procedure relies on a comparison between the measured spectrum and templates for quasars, galaxies and stars. Each class of templates is fitted in a given range of redshift and all the fits are ordered by increasing reduced χ^2. The overall best fit is the fit with the lowest reduced χ^2 and is called *pipeline redshift*. The eBOSS identification procedure is shown in figure 2.14 where five identification flags are used : three flags depending on the classification (star, galaxy or quasar), the pipeline redshift and the ZWARNING flag (when set to 0 it means the classification and redshift are reliable). The spectra of objects with a 'ZWARNING>0' flag and no star identification are visually-inspected.

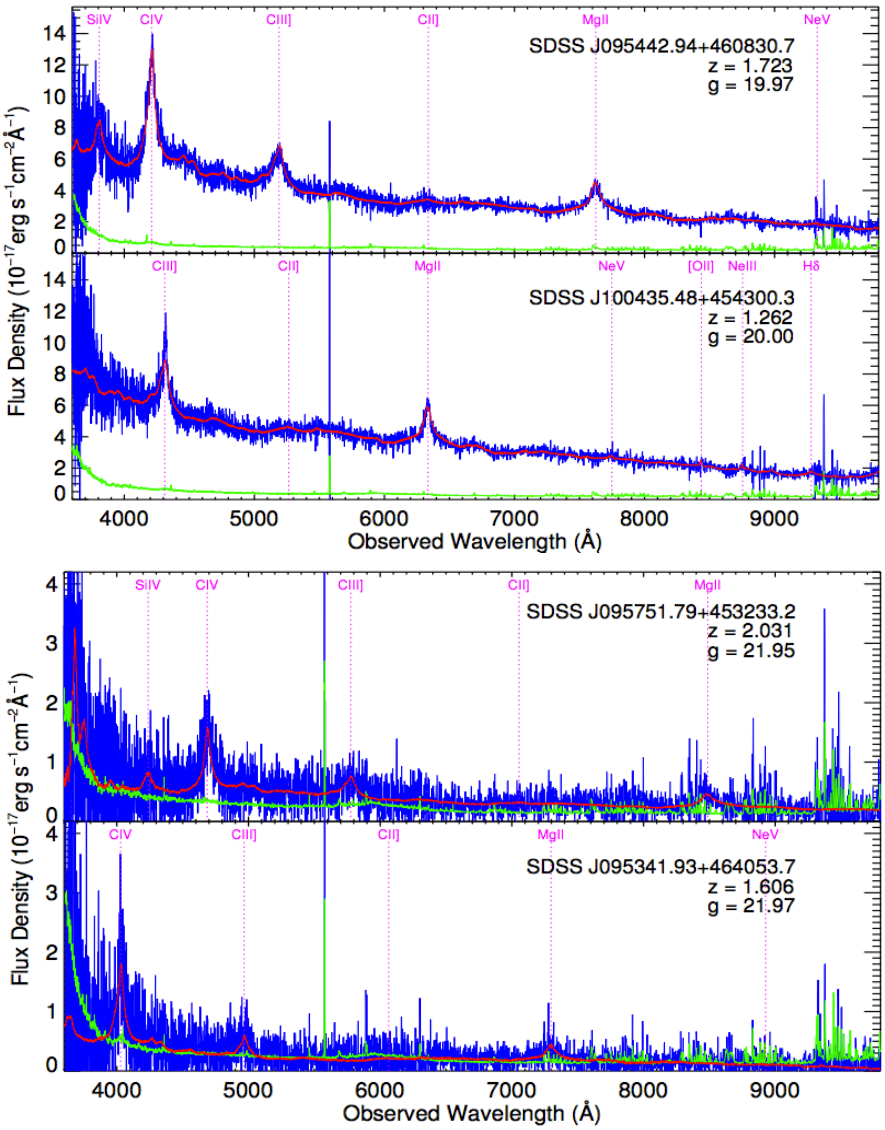


FIGURE 2.13 – eBOSS quasar spectrum for $g = 19$ (top panel) and $g = 22$ (bottom panel) in blue, the red curve is the pipeline model to fit the spectrum and the green component is the sky background whose contribution is removed by the pipeline. As we go to the limit of detection, the quality of the spectra decreases with more noise. From [47].

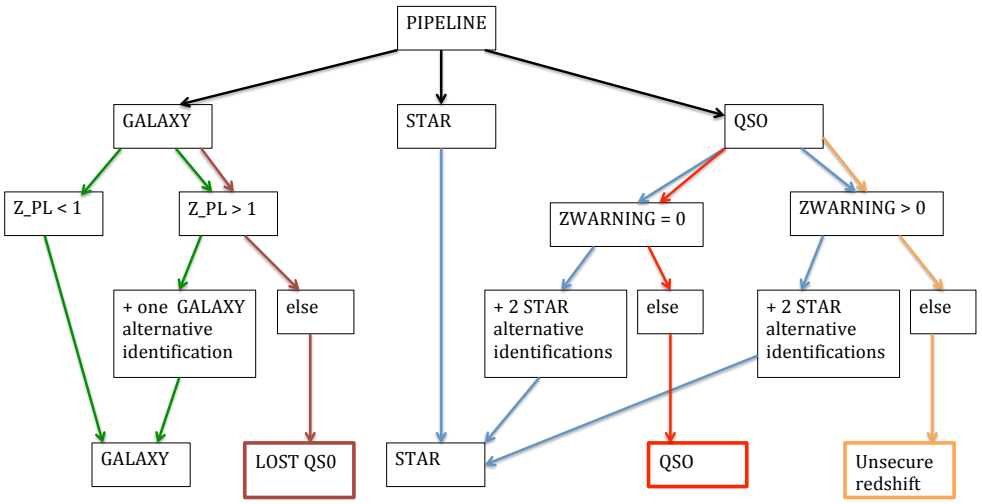


FIGURE 2.14 – eBOSS automatic scheme based on BOSS pipeline using 5 identification flags which include the pipeline response to different templates, the pipeline redshifts and the quality of their estimation (ZWARNING)

2.4.3.2 Visual inspection

Quasars have a broad variety of spectra characteristics that make their automatic classification difficult. To overcome this problem, members from the French Participation Group (FPG) of SDSS have visually inspected every quasar spectrum observed by BOSS. This visual inspection allows a reliable identification as well as a precise determination of the redshift. A full description of this inspection is given in [58]. The visual inspection for BOSS allowed an estimation of the efficiency of the automatic classification of the spectra. On the one hand, it showed that the star sample was very pure with less than 0.1% of the objects classified as stars being quasars in reality. On the other hand, the high redshift ($z > 2.15$) securely-identified quasar sample had more contamination with 0.5% of the objects classified as quasars being either stars or lower-redshift quasars. This contamination was much higher for the low-redshift ($z < 2.15$) quasar sample where it reached 5%. In addition some of the quasars identified by the pipeline were flagged as unsecure detections (ZWARNING) and required a visual inspection to confirm or correct these identifications. This was the case for 8% of the $z > 2.2$ quasars and 25% of the $z < 2.2$ quasars. Almost half of the latter were indeed visually identified as stars and not low redshift quasars. While this was not a concern for BOSS since only high redshift quasars were used in the analyses, this high contamination required a revision of the pipeline for application to eBOSS where both low and high redshift quasars are targeted. Thanks to the visual inspection, it was also possible to check the efficiency of the BOSS automatic redshift determination of quasars [59]. It showed that only 0.3% of the quasars had an error greater than 0.1 ($\Delta z > 0.1$). Such errors mainly occured for quasars with a redshift $z < 2.0$ with a non visible Lyman-α emission line.

The performance of the new pipeline for eBOSS has been tested on SEQUELS observations, the results will be presented in section 2.5. However, despite the improvements on the pipeline, we will also see that a precise determination of the redshift as required for clustering analysis using the eBOSS quasars between $0.8 \leq z \leq 2.2$ remains challenging.

2.5 Redshift determination for quasars

The measurement of distances is a key problem in cosmology and doing so with accuracy is the central path to study the cosmic expansion history and the late-time acceleration of the expansion. Redshift determination proceeds from the analysis of the spectrum of the candidates and it has been shown to be a challenging problem for quasars. Indeed, quasar spectra contain broad emission lines due to the rotating gas located around the central black hole that are subject to matter outflows around the accretion disk. These astrophysical processes frequently give rise to systematic offsets when measuring redshifts. Cosmological measurements that are inferred from the analysis of the clustering of quasars can thus be affected by systematics related to redshift estimate and resolution.

Relation between velocity and redshift shifts

Relative motion in the LOS shifts the lines in the spectrum of any source. One can relate the fractional displacement $d\lambda$ λ in wavelength to the radial velocity by :

$$\frac{v}{c} = \frac{d\lambda}{\lambda} \tag{2.2}$$

And given the definition of the redshift by equation 1.7, any velocity shift will be associated with a shift in redshift as follows

$$\frac{\Delta v}{c} = \frac{\Delta z}{1+z} \tag{2.3}$$

Astrophysists are used to refer to velocity shifts while using redshift seems more natural with spectroscopic data. In what follows, we will use both. We also define a redshift resolution in h^{-1}Mpc to be used in the modelling of the two-point correlation function :

$$\sigma_z = \frac{c\Delta z}{H(z)} \tag{2.4}$$

2.5.1 Quasars redshift estimates

The number and complexity of physical processes that can affect the spectrum of a quasar make it difficult to precisely and accurately disentangle systemic redshift (i.e., as a meaningful indicator of distance) from measured redshift [e.g., 60]. There are two common ways of measuring quasar redshifts from spectroscopic surveys, using either templates or directly the location of individual lines.

Template-based redshift As explained in section 2.4.3.1, the pipeline redshift is determined from the best match between the spectrum and a quasar template derived from a PCA. This template-fitting approach is, by construction, an average representation of the quasar population and does not catch the spectral variations in each individual

spectrum. Another template-based approach has been developed by [58] who derived new PCA components based on two steps :

— A first PCA is performed on a reference sample from SDSS-DR7 [3] using the redshifts provided by [60]. The first four principal components form a basis which is used to fit the MgII emission line of each quasar in the same reference sample.

— Only the spectra where both CIV and MgII are well-detected are kept to compute the new PCA components to be applied to the whole spectrum and optimized to reproduce the quasar spectrum between 1,410 and 2,900 Angstroms in the quasar rest frame where most of the prominent emission lines are covered, in particular CIV and MgII. Table 2.1 summarizes the rest frame wavelengths of the most characteristic lines. A set of four principal components is used to fit the whole spectrum and are dubbed 'PCA redshifts'. A set of five components is used to fit the emission line individually.

Line-based redshift Alternatively, one can measure the observed wavelengths of individual emission lines in the quasar spectrum and then derive line-based redshift estimates like [15] did for the DR14Q catalog. Quasar emission lines are often shifted from the systemic velocity due to various dynamical and radiative processes in the line-emitting region. The level of these velocity shifts depends on the line species and on quasar properties. [61] studied the velocity shifts for various broad and narrow emissioon lines relative to systemic redshift using a sample of 849 quasars from the SDSS-RM program [62]. They were able to measure stellar absorption lines by the host galaxy thanks to the high signal-to-noise ratio of the coadded spectra (from 32 epochs) of individual quasars. These stellar features correspond to absorption in the host galaxy which is not affected by the dynamical and radiative processes associated with the accretion disk around the SMBH and hence provide reliable systemic velocity measurements. At this stage, it is worth defining different kinds of uncertainties in line-based redshifts, following the vocabulary used in [61]

— systematic shift of emission lines w.r.t systemic redshift

— intrinsic scatter in redshift measurements (w.r.t systemic redshift) that can be estimated using individual emission lines

— statistical uncertainty related to the uncertainty in line centers due to the line detectability (weaker lines are presumably more difficult to measure precisely than stronger lines)

The total redshift uncertainty σ_v is defined by :

$$\sigma_v^2 = \sigma_{v,\mathrm{stat}}^2 + \sigma_{v,\mathrm{intr}}^2 \tag{2.5}$$

where $\sigma_{v,\mathrm{stat}}$ is the statistical uncertainty and $\sigma_{v,\mathrm{intr}}$ comes from the intrinsic scatter w.r.t. systemic redshifts. The line with the smallest overall uncertainty provides the best estimate of the systemic redshift. [61] determined empirical guidelines to measure quasar redshifts based on emission lines taking into account the first two kinds of uncertainty. After removing the constant systematic shift and assuming that the measurement uncertainties in the line centers are negligible, the following non-exhaustive list is ordered based on an increasing order of the intrinsic scatter w.r.t systemic redshift :

— Ca II stellar absorption line is assumed to provide the most reliable systemic redshift and thus it has a negligible intrinsic scatter

— [O II] low-ionization narrow emission line with an intrinsic scatter of $\sim 50\,\mathrm{km.s^{-1}}$ and no discernible luminosity dependence

TABLE 2.1 – Wavelengths in rest frame for the most important lines in the eBOSS quasar spectra.

line	type	wavelength in rest frame (Angstroms)
Lyman-α	absorption	1215.67
Carbon IV	emission	1548.2049
	emission	1550.77845
Carbon III	emission	1908.734
Magnesium II	emission	2796.3511
	emission	2803.5324

— [O III] high-ionization narrow emission line has an intrinsic uncertainty of about $50\,\mathrm{km.s^{-1}}$ and a marginal luminosity dependence

— MgII low-ionization broad emission line with an intrinsic uncertainty of $\sim 200\,\mathrm{km.s^{-1}}$ and negligible luminosity dependence

— CIV high-ionization broad emission line has an average blueshift of about $400\,\mathrm{km.s^{-1}}$ with a strong dependence with luminosity

DR14 quasar redshifts The DR14 quasar catalog [15] includes all SDSS-IV/eBOSS objects that were spectroscopically targeted as quasar candidates. The reported redshift estimates are based on the following methods :

— 'z_{PL}' : the SDSS quasar pipeline redshifts which are based on a Principal Component Analysis (PCA) using galaxy, star and quasar templates to fit a linear combination of four eigenspectra to each spectrum [56]. Template-based redshifts are expected to be more stable since they use information from the full spectrum, but at $z \sim 1.5$, the CIV emission line enters the observed spectral range and drives the fit, which has an impact on the redshift acccuracy.

— 'z_{PCA}' : For objects identified as quasars, the redshift is measured using a dedicated PCA of the entire quasar spectrum, and the five principal components are calibrated using the MgII emission as a reference. This approach allows a redshift determination for faint quasars at $z \simeq 2$ when the MgII line approaches the red limit of the SDSS-IV spectral coverage and is not clearly detected.

— 'z_{MgII}' : For objects identified as quasars, MgII-based redshifts are deduced from the location of the maximum of the MgII emission line using the dedicated PCA presented developed in [58].

— 'z_{VI}' : Redshift from visual inspection. For SDSS-III/BOSS, all quasar targets have been visually inspected ; this is not the case for SDSS-IV eBOSS, where only the objects that the automated procedure considers as badly identified lead to a visual inspection [for more details, see Section 3.3 of 15].

2.5.2 Requirements for clustering analysis

With these guidelines in mind, the *Overview* paper [12] determined the requirements in terms of redshift estimate performances to provide the most precise and least biased cosmological measurements from the analysis of the anisotropic clustering of the eBOSS CORE quasar sample between redshifts 0.8 and 2.2. To establish these requirements, they used observations from SEQUELS, the eBOSS pilot program :

Requirement 1 : Rate of catastrophic redshifts *"To minimize the impact of catastrophic erros, we require that the spectroscopic pipeline classifies spectra with fewer than 1% catastrophic errors, where the redshifts are not known to be in error".*
A catastrophic failure corresponds to a redshift estimation that differs from a visual inspection-based redshift by more than $3000\,\mathrm{km.s^{-1}}$. Redshift errors can potentially bias the clustering signal by assigning wrong redshifts because of line confusion or contamination by sky lines. In order to validate the pipeline classification and compare the redshift estimates, all BOSS and SEQUELS quasar spectra have been visually inspected. Several criteria based on pipeline flags are used to determine whether or not an object has been well-identified. The results provided in the *Overview* paper were obtained from visual inspections of all SEQUELS plates observed during BOSS. At this early stage, the spectroscopic pipeline met *Requirement 1* according to which $< 1\%$ of objects are given catastrophically incorrect redshifts. In the DR14 quasar sample, the reference redshift is taken to be the pipeline redshift, exept for visually-inspected quasars when $z_{\mathrm{PL}} - z_{\mathrm{VI}} > 3000\,\mathrm{km.s^{-1}}$, we take z_{VI}. Such a redshift is referred to as 'z' and meets the $< 1\%$ catastrophic redshift requirement. It corresponds to the *reference redshift* for this analysis.

Requirement 2 : Redshift resolution *"We require that the data reduction pipeline provides a combined precision and accuracy of $\sigma_v < 300 km.s^{-1}$ RMS at all redshifts, where σ_v is defined as $cz/(1+z)$. To acknowledge the difficulty in redshift accuracy for the high redshift region of the clustering quasar sample, we relax the requirement to $[300 + 400(z - 1.5)]|, \mathrm{km.s^{-1}}$ for objects at z>1.5.*
Using repeated observations of the same spectrum, [56] found that the RMS scatter of BOSS pipeline redshift estimates is a factor of 2 higher than the reported statistical uncertainty provided by the pipeline. Taking into account this correction, [12] did again the analysis using SEQUELS observations and compared the statistical uncertainty of pipeline and MgII-based redshifts. Their results are shown in Figure 2.15 where they found that the typical statistical precision is about $\sim 100\ km.s^{-1}$ for template-based redshifts and varies between ~ 100 and $\sim 300\ km.s^{-1}$ for MgII-based redshifts depending on the luminosity and on the redshift range. Mg-II based redshifts have been obtained from a sample of 472 quasars in the eBOSS redshift range from the SDSS-RM program with spectra that present a homogeneous S/N distribution corresponding to twice the normal BOSS and eBOSS exposure depth. The redshift and g-band magnitude are comparable with the eBOSS sample. The behavior of the statistical uncertainty for MgII-based redshift wih luminosity is expected since the S/N decreases for fainter objects which makes the MgII line more difficult to be detected. The behavior with redshift for MgII-based redshifts can also be explained : at $z \sim 1.2$, the statistical uncertainty increases as the MgII line moves from the blue to the red arm of the eBOSS spectrograph, and $z \geq 2$, it lies in the red part of eBOSS spectra where the substraction of sky lines is noisier. This comparison thus explains why template-based redshift are often considered as more stable.

However, one must take into the intrinsic uncertainty w.r.t systemic redshift that yields velocity shifts in the measurements of line due to outflows. Given the eBOSS redshift range and typical median signal-to-noise ratio (~ 3 per SDSS pixel compared to ~ 30 for the study led in [61]), the MgII line is present all across the sample between $0.8 \leq z \leq 2$. Moreover, following the guidelines presented in the previous section, the MgII feature is the quasar broad emission line that has the smallest velocity shift (~ 200 km s^{-1}) because it is a lower ionization species that presumably lies at a larger distance from the central black hole. Therefore, it provides the redshift estimate with the smallest intrinsic uncertainty that is available for the eBOSS sample, although its statistical precision is a bit

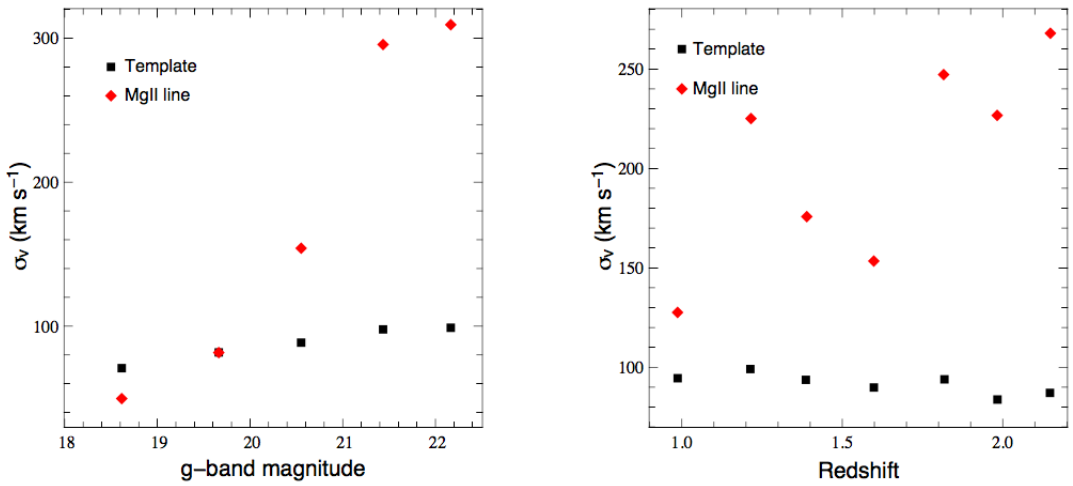


FIGURE 2.15 – Statistical uncertainty on template-based (black squares) and MgII-based (red diamonds) redshift estimates as a function of g-band magnitude (left panel) and of redshift (right panel). From [12].

degraded compared to template-based redshift, particularly for low signal-to-noise lines. In order to estimate the total uncertainty of pipeline redshifts, [12] measured the redshift evolution of the scatter between MgII-based and template-based redshifts. It corresponds to the black points in the top panel of figure 2.16 while the requirements are represented by the red curve.

To quantify the intrinsic scatter of template-based redshifts compared to MgII-based redshift in the DR14 quasar sample, a subsample of quasars is selected where all quasars have the three redshift estimates : $z_{\rm PL}$, $z_{\rm MgII}$ $z_{\rm PCA}$. Figure 2.17 shows the distributions of $\Delta v = \Delta z \cdot c/(1+z)$, for the difference of redshift estimates : $\Delta z = z_{\rm MgII} - z$, $\Delta z = z_{\rm PCA} - z$ and $\Delta z = z_{\rm PCA} - z_{\rm MgII}$ for the two redshift bins in our range of interest. We compare the discrepancies to a Gaussian distribution of width given by the survey requirements (SRD, [12]) where the redshift resolution is showed to be :

$$\sigma_v^{\rm SRD}(z) = 300 \text{ km } s^{-1} \qquad z < 1.5 \qquad (2.6)$$

$$\sigma_v^{\rm SRD}(z) = 400 \times (z - 1.5) + 300 \text{ km } s^{-1} \qquad z > 1.5 \qquad (2.7)$$

The most important feature is that the distributions present large non-Gaussian tails that extend to 3000 km s^{-1}. The distributions involving $z_{\rm MgII} - z$ (green) and $z_{\rm PCA} - z_{\rm MgII}$ (blue) are centered at zero offset (because of the calibration mentioned above) and are mostly symmetric. The distribution obtained for $z_{\rm PCA} - z$ (red) is asymmetric, suggesting that for the special catalogs which mix $z_{\rm PCA}$ and z, there could be systematic shifts in the separation of quasars. The bottom panel of figure 2.17 presents the evolution of the standard deviations of the above distributions as a function of redshift compared to the SRD. When considering only quasars with $|\Delta v| < 1000\,{\rm km.s^{-1}}$ (dashed lines) in the calculation of the standard deviation, our result agrees with the SRD and with the results obtained in [12]. When allowing larger values of $|\Delta v| < 3000\,{\rm km.s^{-1}}$ (solid lines), the standard deviation increases as expected given the shape of the distributions. These results lead to a resolution which is slightly larger than the SRD.

In what follows, the 'z' redshift will be taken as the reference, and in Section 5 the results obtained using this estimate will be compared to the results performed under the

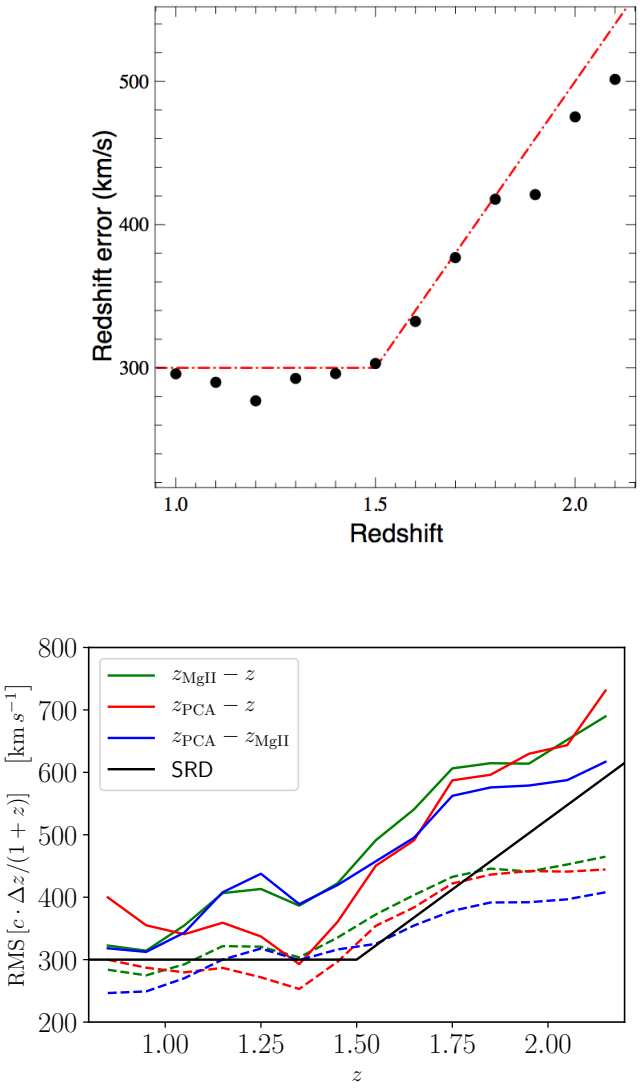


FIGURE 2.16 – Top panel : Redshift evolution of the total scatter of template-based redshifts compared to MgII-based redshifts for a sample of 472 quasars from the SDSS-RM program. This scatter includes the statistical uncertainty and the intrinsic scatter w.r.t. systemic redshift. From [12]. Bottom panel : RMS of the scatter of $\Delta v = \Delta z \cdot c/(1+z)$ for different redshift estimates as a function of redshift, compared to the survey requirements (black solid line). Solid lines (resp. dashed lines) are obtained requiring that $|\Delta v| < 3000 \mathrm{km.s}^{-1}$ (resp. $|\Delta v| < 1000 \mathrm{km.s}^{-1}$).

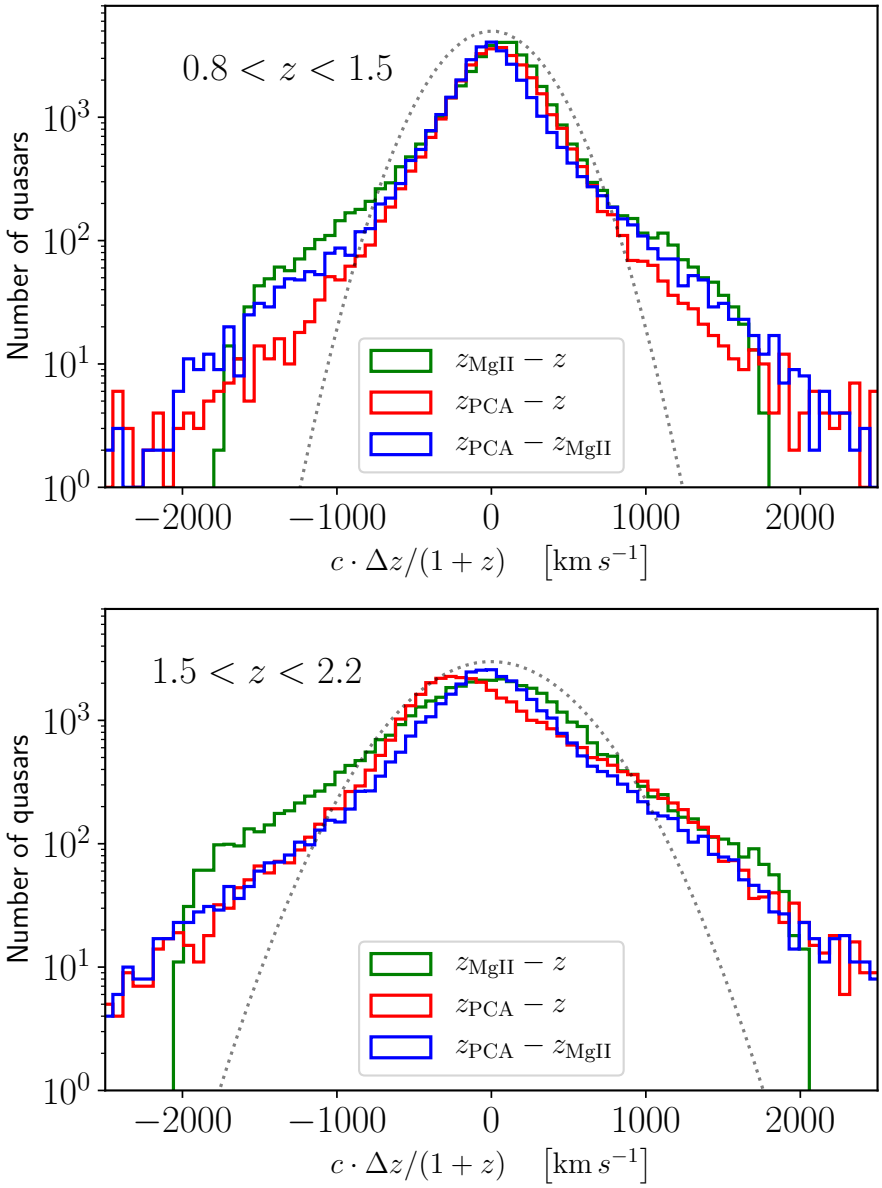


FIGURE 2.17 – Physical distributions (solid lines) of $\Delta v = \Delta z \cdot c/(1+z)$ between different redshift estimates for two redshift bins in our redshift range. The dotted line shows a Gaussian distribution of width given by the survey requirements (see text). The most important feature is that the observed distributions present large non-Gaussian tails that extend to 3000 km s^{-1}. At low redshifts (upper panel), the distributions are mostly symmetric although minor shifts can be observed. At high redshifts (lower panel), the distribution obtained for $z_{\mathrm{PCA}} - z$ (red) is asymmetric, and could yield systematic shifts in the separation of quasars

same conditions but with special catalogs where the redshift is taken to be $z_{\rm MgII}$ (resp. $z_{\rm PCA}$) whenever it is available and z otherwise, such that these catalogs contain the same objects. In Section 4.2, we will also demonstrate that the redshift resolution has a large impact on the clustering signal, especially at scales below 40 h^{-1}Mpc, and that the impact can be measured by fitting the observed clustering. In particular, we will investigate the impact of the redshift resolution on the RSD modeling and on its ability to recover the cosmological parameters both in terms of shape and RMS of the redshift uncertainties distribution.

2.6 eBOSS quasar catalogs for clustering analysis

The creation of the large-scale structure catalogs for clustering analysis involves a number of steps from the list of targets to be observed for spectroscopic observations in addition to the spectroscopic pipeline information about classification and redshift determination. The first LSS catalog was outlined in [63] using the SDSS-III BOSS DR9 data [64] when the survey was approximately one third complete. Then, further refinements of the catalog creation algorithm for the analysis of the second public data release of BOSS, DR10 [65, 66] led to the creation of a new code, called MKSAMPLE. This section is heavily influenced by the description of the large-scale structure catalogs in [67] for the third public SDSS data release DR12 and in [15, 16] for the eBOSS DR14 quasar sample.

In any realistic survey, lots of reasons can lead to variations in the expected number density, such as the survey geometry, inhomogeneities in the observational conditions or instrumental limitations so that the definition of the density contrast (also called the overdensity field or perturbation density field) given in equation 6.1 is affected. It is essential here to understand that a spectroscopic survey does not see all galaxies that lie in a given direction. This is represented by the *survey selection function* $\phi(\mathbf{x})$, which gives the probability for an object at distance $\mathbf{x}$ to be observed by the survey. Thus, the basic assumption of the two-point clustering analysis is that the expected number density, $n_{\rm exp}$, for a constant underlying number density, $\bar{n}$; is given by the survey selection function :

$$n_{\rm exp} = \phi(\mathbf{x})\bar{n} \tag{2.8}$$

It is usually assumed that the survey selection function can be written as the product of an angular selection function $\phi_{\rm ang}$ and a radial part $\phi_{\rm rad}$. It means that we assume the survey depth and hence redshift distributions of selected objects do not vary significantly with position on the sky. It is a good approximation for the eBOSS quasar sample as the redshift distribution of quasars does not depend strongly on luminosity [47] and we will see that the main variations with the survey depth can be corrected for by applying weights. In contrast, the redshift distribution of BOSS galaxies has a slight dependence with stellar density depended on the surface brightness of the galaxies. (the lowest surface brightness galaxies in this sample are at high redshift) so the weights are defined as a function of the local stellar density and the surface brightness of the galaxy in order to remove any redshift dependence [53]. The angular selection function consists in defining a survey mask that corrects for selection effects due to the observational strategy and which will therefore reduce the effective footprint used for clustering analysis. The radial selection function requires to understand the spectroscopic procedure and redshift measurements to ensure that the observed redshift distribution is truly representative of the sample we are analyzing.

2.6.1 The angular selection function

In order to characterize the angular selection function $\phi_{\rm ang}$ of the eBOSS survey, the eBOSS footprint is split into *sectors* that correspond to the assembly of spherical polygons defined by a unique intersection of spectroscopic tiles. Spectroscopic tiles are the output of the tiling algorithm as explained in Section 2.4.2.2. Spherical polygons are used to represent the boundaries of the survey geometry which include the circular fields defined by spectroscopic tiles and small regions to be removed. These small regions are masked because galaxies could not have been observed, such as the centerposts of each spectroscopic tile (where no targets coinciding with the centerpost can be observed) or the surroundings of bright objects or regions with missing or bad photometric information. Despite the small angular size of each individual 'masked' region, they are not randomly distributed across the sky and sum to a non-negligible area. Thus, they are excluded from any analysis by the use of veto masks. For the eBOSS DR14 sample, the bad photometric fields exclude approximately 5% of the area, the bright stars mask 1.8%, the bright objects mask 0.05% and the centerposts mask less than 0.01% of the area [16].

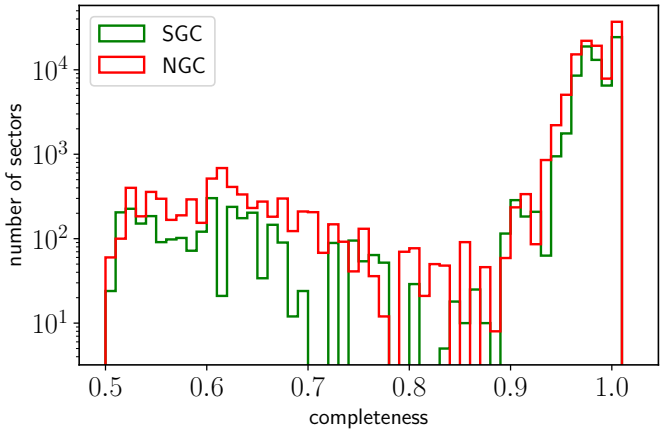


FIGURE 2.18 – Distribution of the completeness $C_{\rm eBOSS}$ per sector. Sectors with completeness smaller than 0.9 correspond to overlaping plates regions where only one plate has currently been measured. Objects with $C_{\rm eBOSS} < 0.5$ are not considered for clustering analysis.

We use the MANGLE software [68] to construct the survey mask from spherical polygons which form the base unit for the geometrical decomposition of the sky. For each sector, an observational completeness $C_{\rm eBOSS}$ is calculated from $\rm N_{target}$, the number of imaging quasar targets selected, $\rm N_{fiber}$, the number of targets that actually received a fiber after the tiling algorithm is applied, $\rm N_{cp}$, the number of targets that were in collision within the 62″ exclusion radius around each target and thus did not receive a fiber, and $\rm N_{known}$ the number of targets that are confirmed quasars measured by previous surveys at the time of tiling (and called *known*). The traditional way of accounting for the missing quasars is to up-weight the nearest quasar (more details are provided in section 2.6.2) so that the observational completeness is defined as :

$$C_{\rm eBOSS} = \frac{\rm N_{fiber} + N_{cp}}{\rm N_{target} - N_{known}}. \tag{2.9}$$

Known targets are 100% complete, since they have already been observed and they are sub-sampled to match the value of $C_{\rm eBOSS}$ in each sector. Low completeness sectors are

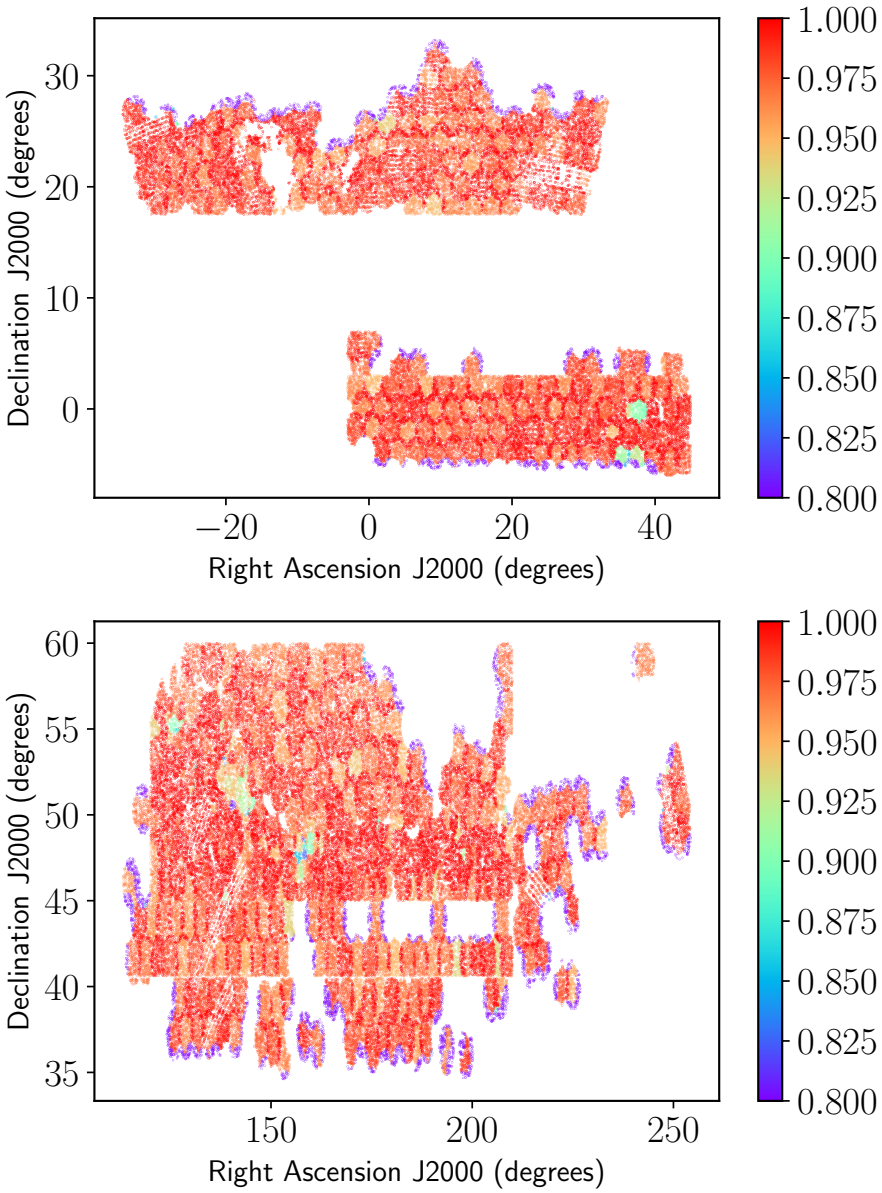


Figure 2.19 – Footprint of the DR14Q catalog used for this analysis. The upper (lower) panel displays the SGC (NGC) resulting in a total effective area of 2112.9 deg^2. Each object is color-coded according to the completeness of the sector to which it belongs (object in purple have completeness between 0.5 and 0.8)

due to overlaping plates for which some plates have not been observed yet. In particular, the mean completeness for two overlaping plates is $C_{\rm eBOSS} \sim 0.5$ so we keep all sectors with $C_{\rm eBOSS} > 0.5$ to guarantee good observational conditions across the footprint. The average completeness of the remaining sectors is high : $\sim 95\%$ and $\sim 96\%$ in the North Galactic Cap (NGC) and South Galactic Cap (SGC), respectively. The distribution of the completeness per sector for our survey is shown in Figure 2.18. However it is not 100% in all sectors because the CORE quasar targets do not get the highest priority and because of combinatorial requirements in the tiling algorithm. The footprint of spectroscopically-observed objects is shown in Figure 2.19, color-coded according to the completeness. The large gap in the SGC at RA ~ -10 deg and Dec $\sim 22,$ deg is due to the mask for Galactic extinction and the horizontal striped patterns are due to the photometric bad fields or poor seing in the SDSS imaging. The other masks are too small to be distinguishable.

2.6.2 Accounting for observational artefacts

As mentioned above, the number density of quasars observed in eBOSS, as in any realistic survey, suffers from survey systematics that arise because of inhomogeneities in the targeting and observing strategies. In order to mimize the impact of observational artefacts on our estimate of the true galaxy overdensity field, we apply weights to the targeted quasars. This approach has first been used for the clustering analysis of BOSS DR9 galaxies in [66], the weights have thus been developed in [67] for BOSS DR12 and we use a similar approach for the analysis with the eBOSS DR14 quasar sample. The corrective weights to correct for the survey incompleteness are :

— **Photometric weight** : [69] demonstrated that the density of stars had a significant effect on the observed density of BOSS galaxies, which could introduce spurious fluctuations in the target density field. They showed that this effect could be corrected for using a weight that mimized these fluctuations as a function of stellar density as long as the target density fluctuations are less than 15%. The determination of this weight has been developed to take into account other dependencies with photometric conditions such as the seeing, Galactic extinction, airmass and sky background [53]. The photometric weight is written $w_{\rm photo}$.

— **Fiber-collisions weight** : The angular size of a spectroscopic fiber is 62" (which corresponds to $0.54h^{-1}$Mpc at $z = 1.5$) and prevents observing two quasars within a radius of 62" with a single observational plate. When the other target is also a quasar, those collisions are called 'fiber' or 'close pair' collisions. When the target is of a different type, the collisions are missed. A rough estimation of the number of collisions between two quasars CORE is given by $(62''/3600'')^2\,\mathrm{deg}^2 \times 115\mathrm{targets}/\,\mathrm{deg}^2 = 0.034$, i.e. about 30% of the collisions are between two CORE quasars. The targets that were not assigned a spectroscopic fiber due to fiber collisions are not a random subsample of the full target sample, they are more likely to occur in overdense regions. The traditional way of correcting for that is to transfer the weight of the lost target to the nearest neighbour of the same target class with a valid redshift (for limitations of this approach, see [70]). This weight is written $w_{\rm cp}$.

— **Redshift failures weight** : Not all observations yield a valid redshift and [53, 34] showed that it does not happen randomly on a tile but depend on the fiber used. As for fiber-collisions, we can transfer the weight of the lost target to the nearest neighbour of the same target class and with a good redshift. This weight is written $w_{\rm noz}$.

— **FKP weight** : We also apply weights to optimize our clustering measurements w.r.t. to shot-noise and cosmic variance. In order to compensate for the different signal-to-noise ratio produced by these variations, [71] showed that an optimal compromise is to assign weights according to the measured mean density of quasars at that redshift, $\bar{n}(z)$. We thus use the FKP weights defined as $w_{\rm FKP} = 1/[1 + \bar{n}(z)P_0]$ where P_0 is the amplitude of the power spectrum at the k scale at which the BAO signature in the DR14 quasar sample has the highest signal. It corresponds to $k \sim 0.14h.{\rm Mpc}^{-1}$ which gives $P_0 = 6 \times 10^3 [h^{-1}.{\rm Mpc}]^3$. Including this weight equalises the contribution of every redshift interval that are considered in the calculation of the correlation function. However, the statistical gain brought by the FKP-weight is small because the number density of eBOSS quasars is low and nearly constant, as shown in the left panel of Figure 2.20, so the value of the weight varies by less than 10%.

Therefore, the total weighting scheme and the effective number of quasars in the sample are defined by :

$$\mathcal{W}_{\rm tot} = w_{\rm FKP} \cdot w_{\rm photo} \cdot (w_{\rm cp} + w_{\rm noz} - 1) \tag{2.10}$$

$$N_{\rm quasars,eff} = \sum_i \mathcal{W}_{\rm tot,i} \tag{2.11}$$

This weighting scheme is the standard scheme that has been applied to the analysis of the BOSS galaxies and to the BAO analysis of the DR14 quasar sample [16]. Section 4.3 will present in more details the effect of the weights on the clustering and will describe the improved weighting scheme that has been used for the RSD analysis of the DR14 quasar sample in configuration space [17], which is the subject of this thesis.

2.6.3 The radial selection function

The angular target density can be converted into a redshift probability distribution defined in [67] by :

$$p(z_j, z_j + dz)dz \propto \frac{\sum_{z_i \in [z_j, z_j+dz]} \mathcal{W}_{\rm tot,i}}{\sum_i \mathcal{W}_{\rm tot,i}} \tag{2.12}$$

where the numerator corresponds to the effective number of quasars in a given redshift bin and the denominator is the total effective number of quasars in the whole redshift range. The inclusion of the $w_{\rm photo}$ in the definition of $\mathcal{W}_{\rm tot}$ given by equation 6.5 accounts for any impact of the angular systematics on the redshift distribution. We then use the fiducial cosmology mentioned previously to determine the number of targets per $[h^{-1}.{\rm Mpc}]^3$. The observed quasar density is shown in the left panel of Figure 2.20 where one can see small differences between the NGC (blue) and SGC (yellow). Such differences have already been reported in [72] and are related to regions in the sky with different sensitivities to the target selection. The redshift distribution of the CORE quasars in the DR14 catalog is presented in the right panel of Figure 2.20. The orange histogram corresponds to the distribution of the known quasars at the start of eBOSS data taking. Over 75% of the new redshifts were obtained during the eBOSS program.

The DR14 quasar catalog [15] contains 158,757 objects between $0.8 \leq z \leq 2.2$ identified as quasars by the pipeline. 93.6% of them have a secure idenfication with a valid pipeline redshift, they represent the 148,659 quasars used in this analysis. About 13% quasars have been visually inspected (more details on the visual inspection procedure can be found in [15]). Regarding the remaining 10,090 objects, 5,188 did not receive a spectroscopic observation because of fiber-collisions but 1,015 (20%) have been resolved thanks to plate overlap. 4,910 objects were securely classified by the pipeline but their redshifts were

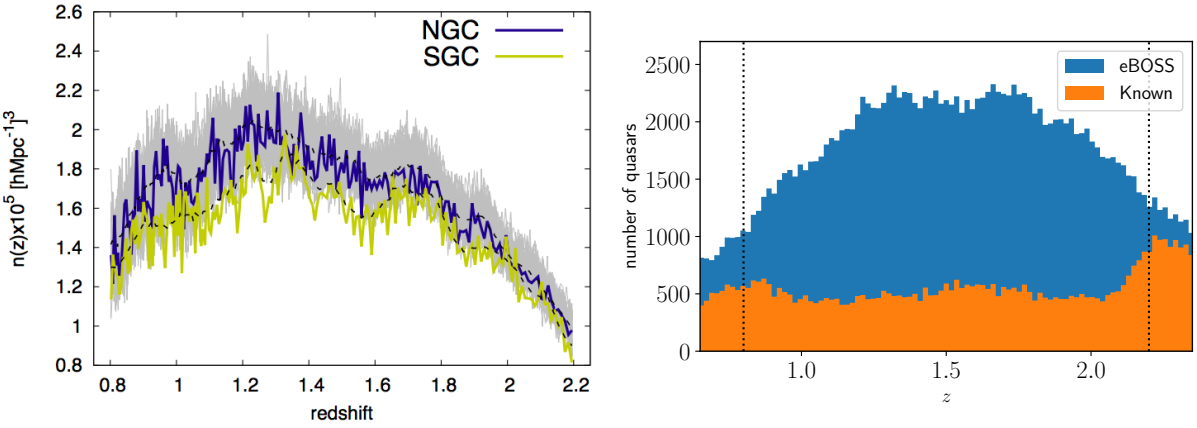


FIGURE 2.20 – Left panel : Number density distribution as a function of redshift for the NGC and the SGC. From [73]. Right panel : Redshift distribution of the objects in the DR14 catalog corresponding to the CORE sample. The orange histogram shows the known quasars at the start of eBOSS data taking. Objects in $0.8 \leq z \leq 2.2$ are kept for this analysis.

TABLE 2.2 – Number of quasars with $0.8 \leq z \leq 2.2$ of the eBOSS CORE sample and effective area for the North Galactic Cap (NGC) and South Galactic Cap (SGC).

	NGC	SGC	Total
N_{quasar} $(0.8 \leq z \leq 2.2)$	89233	59426	148659
Effective area (deg^2)	1214.6	898.3	2112.9

not valid and they were not visually-inspected. Eventually, it means that fiber-collisions and redshift-failures represent ∼3% of the total sample each.
The effective number of objects and area of the sample are given in Table 6.1 and correspond to a maximum density of $n_{\rm obs,max} = 2 \times 10^{-5} h^3 {\rm Mpc}^{-3}$.

We define the effective redshift $z_{\rm eff}$ as $z_{\rm eff} = \sum_i z_i \mathcal{W}_{\rm tot,i} \ / \ \sum_i W_{\rm tot,i} = 1.52$ and we follow the definition of the effective volume covered by the footprint given in [67] :

$$V_{\rm eff} = \sum_i \left(\frac{\bar{n}_i P_0}{1 + \bar{n}_i P_0} \right)^2 \Delta V(z_i) \tag{2.13}$$

where $\Delta V(z_i$ is the volume of the shell at z_i and we use $P_0 = h^{-3}.{\rm Mpc}^3$ at $k = 0.14h{\rm Mpc}$ where the BAO signal is the strongest. The DR14 quasar sample represents an effective volume of 0.246 $\rm Gpc^3$, but the associated comoving volume is about 32 $\rm Gpc^3$. The difference is due to the factor $\bar{n}_i P_0/(1 + \bar{n}_i P_0)$ in the definition of the effective volume. The DR14 quasar sample is also a sparse sample as the number density of quasars is an order of magnitude lower compared to SDSS BOSS galaxies (about $\sim 10^{-4}[h.Mpc]^{-3}$). Indeed $n_{\rm obs} P \ll 1$ for the eBOSS quasar sample which indicates that the sample is shot noise dominated, where $n_{\rm obs}$ is the observed quasar density and P is the amplitude of the power spectrum at the BAO scale ($k = 0.14h.{\rm Mpc}^{-1}$). On the contrary, BOSS galaxies have $n_{\rm obs} P \gg 1$ so that the sample is dominated by the cosmic variance.

2.6.4 Generation of the random catalog

We generate a random catalog of unclustered objects with the detailed angular and redshift selection functions of the data sample in order to account for the complex survey geometry. Together, the data and random catalogs constitute the input for clustering analysis.
The whole survey selection is described by a set of random points which sample the survey volume with a density which is taken to be roughly 40 times higher than the one of quasars to ensure that the survey geometry is well-sampled. Angular positions of objects in the random catalog are generated according to the completeness assigned to each sector and taking into account the veto masks so that the angular distribution of randoms follows the angular selection function of the survey. To take into account the selection function, we must assign redshifts to the random catalog. [53] compared different methods to simulate the radial selection function and showed that the 'shuffled' technique where the redshifts were drawn randomly from the measured redshift distribution provided the smallest bias on the monopole and quadrupole of the correlation function. We thus use this method and assign redshifts to the objects in the random catalog by randomly selecting a redshift in the measured quasar redshift distribution. This procedure ensures that the weighted quasar and random catalogs have exactly the same redshift distribution. Then, each quasar in the random catalog is given a FKP weight $w_{\rm FKP}$ only, as fluctuations in the target density and spectroscopic incompleteness are corrected in the data catalog directly.

2.7 Estimation of the two-point correlation function

2.7.1 Assume a fiducial cosmology

At this stage of the analysis, we have a catalog of spectroscopically-confirmed quasars with their positions in the sky, namely two angular coordinates (right ascension α, declination δ) and the redshift, from which we want to compute the two-point correlation function. First, the observational coordinates are translated into spherical coordinates assuming a cosmological model :

$$r = D_C \tag{2.14}$$
$$\theta = 90 - \delta \tag{2.15}$$
$$\phi = \alpha \tag{2.16}$$

The first equation implies the Hubble constant H_0 and the comoving distance D_C defined by equation 1.23 which implies to know the values for Ω_m and Ω_Λ. Spherical coordinates are then translated into cartesian coordinates x, y and z.
In this work, measured redshift and angular coordinates are converted to comoving coordinates using the fiducial cosmology that was used for the BOSS DR12 analysis [8] and for the eBOSS DR14Q BAO analysis [16], where the universe is assumed to be flat with

$$h = 0.676,\ \Omega_m = 0.31,\ \Omega_\Lambda = 0.69, \Omega_b h^2 = 0.022,\ \sigma_8 = 0.80 \tag{2.17}$$

If the wrong cosmological parameters are used to make this transformation then we can induce anisotropic distortions in the observed correlation function which are similar to the RSD signal [74]. By fitting simultaneously the AP parameters given in equations 6.14 with the growth rate, we can separate both effects and provide a measurement of the growth rate which is independent of the assumption of sphericity.

2.7.2 Define the most appropriate estimator

Methods to estimate the two-point correlation function from the data are based on its definition as an excess probability of finding a galaxy pair which is given by equation 1.43. One counts from the data catalog the number $DD(r)$ of pairs of galaxies with separation $|\mathbf{r}_2 - \mathbf{r}_1| \in r \pm \frac{1}{2}\Delta r$ where Δr is the bin width (for what follows, "separation r" means that the separation falls in this bin). To determine the fraction of excess in this bin, we need to compare the number of pairs in the data catalog to the number of pairs we would get for $\xi = 0$, i.e. for a Poisson process which corresponds to a random distribution. For this, we generate a random catalog by generating points randomly in the survey volume and count the corresponding number $RR(r)$ of random pairs. This approach leads to a simple estimator of the correlation function, $\hat{\xi}(r)$ defined as follows :

$$1 + \hat{\xi}(r) = \frac{DD(r)}{RR(r)} \tag{2.18}$$

This definition assumes that both the data and the random catalogs were generated from an ensemble with the same expectation value of the density field $\langle\rho\rangle$. The problem is that we do not know the true expectation value. We have to estimate it fom the data as N_D/V where N_D is the number of data points (observed galaxies) and V the survey volume. The estimator defined by equation 2.18 assumes an equal number of random points $N_\mathrm{R} = N_\mathrm{D}$. The inherent randomness in both the data and the random catalogs introduces a random error in the estimator, the larger the number of points the smaller the expected random error. The usual strategy is to increase N_R to minimize random errors due to finite number of points. One then needs to scale the estimator with the ratio of total number of possible random pairs $\frac{1}{2}N_\mathrm{R}(N_\mathrm{R} - 1)$ to that of the data pairs $\frac{1}{2}N_\mathrm{D}(N_\mathrm{D} - 1)$. Thus, the estimator becomes :

$$1 + \hat{\xi}(r) = \frac{N_\mathrm{R}(N_\mathrm{R} - 1)DD(r)}{N_\mathrm{D}(N_\mathrm{D} - 1)RR(r)} \tag{2.19}$$

Moreover, since a realistic survey samples just a finite volume of a single realization, the estimation of the correlation function also differs from the true correlation function, because of the survey geometry and of masks which lead to *edge effects* for instance. We can try to minimize these edge effects by using data-random pairs $DR(r)$ where the total number of data-random pair counts is $N_\mathrm{D}N_\mathrm{R}$. One can thus construct different estimators using $DD(r)$, $RR(r)$ and $DR(r)$. [75] compared four estimators of ξ in terms of bias and variance in the limit where the correlations are small ($\xi << 1$) and showed that the one with the miminum variance corresponds to :

$$\hat{\xi}(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)} \tag{2.20}$$

This estimator is now called the *Landy-Szalay* (LS) estimator.

In practice, as there is a remaining invariance w.r.t. rotations around the LOS axis, the two-point statistics are two-dimensional. The second parameter for the correlation function besides the distance s in redshift space is the parameter $\mu = \cos\theta$ where θ is the angle between the LOS and the orientation vector of the pair of tracers under consideration. The definition is illustrated in Figure 2.21 where μ has values in the range 0 to 1 ($\theta \in [0, \pi]$).

In this work, the two-point correlation function $\xi(s, \mu)$ of the data is determined using the LS estimator.

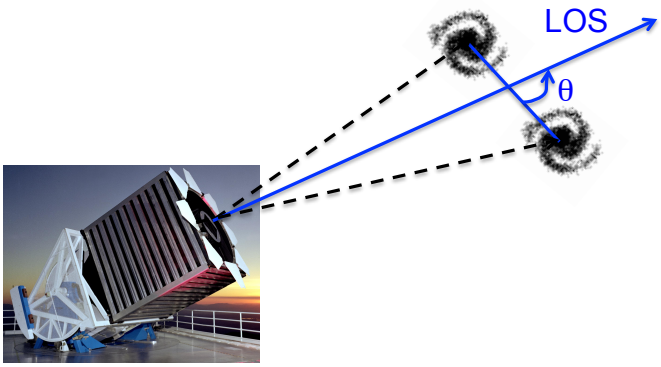


FIGURE 2.21 – The anisotropic clustering is a function of the separation vector between two quasars and the cosine of θ, the angle between the LOS and the orientation vector of the pair.

Bibliographie

[1] A. Finkbeiner, *A Grand and Bold Thing* (2010).

[2] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, *et al.*, *The Sloan Digital Sky Survey : Technical Summary*, **120**, 1579 (2000), doi:10.1086/301513, astro-ph/0006396.

[3] K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, *et al.*, *The Seventh Data Release of the Sloan Digital Sky Survey*, **182**, 543 (2009), doi:10.1088/0067-0049/182/2/543, arXiv:0812.0649.

[4] B. Yanny, C. Rockosi, H. J. Newberg, G. R. Knapp, *et al.*, *SEGUE : A Spectroscopic Survey of 240,000 Stars with g = 14-20*, **137**, 4377 (2009), doi:10.1088/0004-6256/137/5/4377, arXiv:0902.1781 [astro-ph.GA].

[5] J. A. Frieman, B. Bassett, A. Becker, C. Choi, *et al.*, *The Sloan Digital Sky Survey-II Supernova Survey : Technical Summary*, **135**, 338 (2008), doi:10.1088/0004-6256/135/1/338, arXiv:0708.2749.

[6] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, *et al.*, *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies*, **633**, 560 (2005), doi:10.1086/466512.

[7] D. J. Eisenstein, D. H. Weinberg, E. Agol, H. Aihara, *et al.*, *SDSS-III : Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems*, **142**, 72 (2011), doi:10.1088/0004-6256/142/3/72, arXiv:1101.1529.

[8] S. Alam, M. Ata, S. Bailey, F. Beutler, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : cosmological analysis of the DR12 galaxy sample*, **470**, 2617 (2017), doi:10.1093/mnras/stx721, arXiv:1607.03155.

[9] J. E. Bautista, N. G. Busca, J. Guy, J. Rich, *et al.*, *Measurement of baryon acoustic oscillation correlations at z = 2.3 with SDSS DR12 Lyα-Forests*, **603**, A12 (2017), doi:10.1051/0004-6361/201730533, arXiv:1702.00176.

[10] H. du Mas des Bourboux, J.-M. Le Goff, M. Blomqvist, N. G. Busca, *et al.*, *Baryon acoustic oscillations from the complete SDSS-III Lyα-quasar cross-correlation function at $z = 2.4$*, ArXiv e-prints (2017), arXiv:1708.02225.

[11] M. R. Blanton, H. Lin, R. H. Lupton, F. M. Maley, *et al.*, *An Efficient Targeting Strategy for Multiobject Spectrograph Surveys : the Sloan Digital Sky Survey "Tiling" Algorithm*, **125**, 2276 (2003), doi:10.1086/344761, astro-ph/0105535.

[12] K. S. Dawson, J.-P. Kneib, W. J. Percival, S. Alam, *et al.*, *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey : Overview and Early Data*, **151**, 44 (2016), doi:10.3847/0004-6256/151/2/44, arXiv:1508.04473.

[13] G.-B. Zhao, Y. Wang, A. J. Ross, S. Shandera, *et al.*, *The extended Baryon Oscillation Spectroscopic Survey : a cosmological forecast*, **457**, 2377 (2016), doi:10.1093/mnras/stw135.

[14] B. Abolfathi, D. S. Aguado, G. Aguilar, C. Allende Prieto, *et al.*, *The Fourteenth Data Release of the Sloan Digital Sky Survey : First Spectroscopic Data from the extended Baryon Oscillation Sky Survey and from the second phase of the Apache Point Observatory Galactic Evolution Experiment*, ArXiv e-prints (2018), arXiv:1707.09322.

[15] I. Pâris, P. Petitjean, É. Aubourg, A. D. Myers, *et al.*, *The Sloan Digital Sky Survey Quasar Catalog : Fourteenth data release*, **613**, A51 (2018), doi:10.1051/0004-6361/201732445, arXiv:1712.05029.

[16] M. Ata, F. Baumgarten, J. Bautista, F. Beutler, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : First measurement of Baryon Acoustic Oscillations between redshift 0.8 and 2.2*, ArXiv e-prints (2017), arXiv:1705.06373.

[17] P. Zarrouk, E. Burtin, H. Gil-Marín, A. J. Ross, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2*, **477**, 1639 (2018), doi:10.1093/mnras/sty506.

[18] J. E. Gunn, W. A. Siegmund, E. J. Mannery, R. E. Owen, *et al.*, *The 2.5 m Telescope of the Sloan Digital Sky Survey*, **131**, 2332 (2006), doi:10.1086/500975, astro-ph/0602326.

[19] S. A. Smee, J. E. Gunn, A. Uomoto, N. Roe, *et al.*, *The Multi-object, Fiber-fed Spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey*, **146**, 32 (2013), doi:10.1088/0004-6256/146/2/32, arXiv:1208.2233.

[20] J. E. Gunn, M. Carr, C. Rockosi, M. Sekiguchi, *et al.*, *The Sloan Digital Sky Survey Photometric Camera*, **116**, 3040 (1998), doi:10.1086/300645.

[21] W. Baade et R. Minkowski, *Identification of the Radio Sources in Cassiopeia, Cygnus A, and Puppis A.*, **119**, 206 (1954), doi:10.1086/145812.

[22] A. S. Bennett, *The preparation of the revised 3C catalogue of radio sources*, **125**, 75 (1962), doi:10.1093/mnras/125.1.75.

[23] T. A. Matthews et A. R. Sandage, *Optical Identification of 3C 48, 3C 196, and 3C 286 with Stellar Objects.*, **138**, 30 (1963), doi:10.1086/147615.

[24] J. L. Greenstein et M. Schmidt, *The Quasi-Stellar Radio Sources 3C 48 and 3C 273.*, **140**, 1 (1964), doi:10.1086/147889.

[25] E. E. Salpeter, *Accretion of Interstellar Matter by Massive Objects.*, **140**, 796 (1964), doi:10.1086/147973.

[26] D. Lynden-Bell, *Galactic Nuclei as Collapsed Old Quasars*, **223**, 690 (1969), doi:10.1038/223690a0.

[27] C. K. Seyfert, *Nuclear Emission in Spiral Nebulae.*, **97**, 28 (1943), doi:10.1086/144488.

[28] D. Richstone, E. A. Ajhar, R. Bender, G. Bower, *et al.*, *Supermassive black holes and the evolution of galaxies.*, **395**, A14 (1998), astro-ph/9810378.

[29] L. Ferrarese et D. Merritt, *A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies*, **539**, L9 (2000), doi:10.1086/312838, astro-ph/0006053.

[30] K. Gebhardt, R. Bender, G. Bower, A. Dressler, *et al.*, *A Relationship between Nuclear Black Hole Mass and Galaxy Velocity Dispersion*, **539**, L13 (2000), doi:10.1086/312840, astro-ph/0006289.

[31] J. Kormendy et D. Richstone, *Inward Bound—The Search For Supermassive Black Holes In Galactic Nuclei*, **33**, 581 (1995), doi:10.1146/annurev.aa.33.090195.003053.

[32] M. White, A. D. Myers, N. P. Ross, D. J. Schlegel, *et al.*, *The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey*, **424**, 933 (2012), doi:10.1111/j.1365-2966.2012.21251.x, arXiv:1203.5306.

[33] S. Eftekharzadeh, A. D. Myers, M. White, D. H. Weinberg, *et al.*, *Clustering of intermediate redshift quasars using the final SDSS III-BOSS sample*, ArXiv e-prints (2015), arXiv:1507.08380.

[34] P. Laurent, S. Eftekharzadeh, J.-M. Le Goff, A. Myers, *et al.*, *Clustering of quasars in SDSS-IV eBOSS : study of potential systematics and bias determination*, **7**, 017 (2017), doi:10.1088/1475-7516/2017/07/017, arXiv:1705.04718.

[35] D. B. Sanders, B. T. Soifer, J. H. Elias, B. F. Madore, *et al.*, *Ultraluminous infrared galaxies and the origin of quasars*, **325**, 74 (1988), doi:10.1086/165983.

[36] R. G. Carlberg, *Quasar evolution via galaxy mergers*, **350**, 505 (1990), doi:10.1086/168406.

[37] P. F. Hopkins, R. S. Somerville, L. Hernquist, T. J. Cox, *et al.*, *The Relation between Quasar and Merging Galaxy Luminosity Functions and the Merger-driven Star Formation History of the Universe*, **652**, 864 (2006), doi:10.1086/508503, astro-ph/0602290.

[38] C. Conroy et M. White, *A Simple Model for Quasar Demographics*, **762**, 70 (2013), doi:10.1088/0004-637X/762/2/70, arXiv:1208.3198.

[39] M. Fukugita, T. Ichikawa, J. E. Gunn, M. Doi, *et al.*, *The Sloan Digital Sky Survey Photometric System*, **111**, 1748 (1996), doi:10.1086/117915.

[40] D. P. Finkbeiner, E. F. Schlafly, D. J. Schlegel, N. Padmanabhan, *et al.*, *Hypercalibration : A Pan-STARRS1-based Recalibration of the Sloan Digital Sky Survey Photometry*, **822**, 66 (2016), doi:10.3847/0004-637X/822/2/66.

[41] E. L. Wright, P. R. M. Eisenhardt, A. K. Mainzer, M. E. Ressler, *et al.*, *The Wide-field Infrared Survey Explorer (WISE) : Mission Description and Initial On-orbit Performance*, **140**, 1868-1881 (2010), doi:10.1088/0004-6256/140/6/1868, arXiv:1008.0031 [astro-ph.IM].

[42] J. R. Pier, J. A. Munn, R. B. Hindsley, G. S. Hennessy, *et al.*, *Astrometric Calibration of the Sloan Digital Sky Survey*, **125**, 1559 (2003), doi:10.1086/346138.

[43] R. Lupton, J. E. Gunn, Z. Ivezić, G. R. Knapp, *et al.*, in *Astronomical Data Analysis Software and Systems X* (2001), vol. 238, p. 269.

[44] C. Stoughton, J. Adelman, J. T. Annis, J. Hendry, *et al.*, in *Survey and Other Telescope Technologies and Discoveries* (2002), vol. 4836, pp. 339–349.

[45] N. Padmanabhan, D. J. Schlegel, D. P. Finkbeiner, J. C. Barentine, *et al.*, *An Improved Photometric Calibration of the Sloan Digital Sky Survey Imaging Data*, **674**, 1217 (2008), doi:10.1086/524677.

[46] E. F. Schlafly, D. P. Finkbeiner, M. Jurić, E. A. Magnier, *et al.*, *Photometric Calibration of the First 1.5 Years of the Pan-STARRS1 Survey*, **756**, 158 (2012), doi:10.1088/0004-637X/756/2/158, arXiv:1201.2208 [astro-ph.IM].

[47] A. D. Myers, N. Palanque-Delabrouille, A. Prakash, I. Pâris, *et al.*, *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey : Quasar Target Selection*, **221**, 27 (2015), doi:10.1088/0067-0049/221/2/27, arXiv:1508.04472.

[48] A. Sandage, *The Existence of a Major New Constituent of the Universe : the Quasistellar Galaxies.*, **141**, 1560 (1965), doi:10.1086/148245.

[49] P. C. Hewett, C. B. Foltz, et F. H. Chaffee, *The Large Bright Quasar Survey.VI.Quasar Catalog and Survey Parameters*, **109**, 1498 (1995), doi:10.1086/117380.

[50] S. M. Croom, R. J. Smith, B. J. Boyle, T. Shanks, *et al.*, *The 2dF QSO Redshift Survey - XII. The spectroscopic catalogue and luminosity function*, **349**, 1397 (2004), doi:10.1111/j.1365-2966.2004.07619.x.

[51] S. M. Croom, G. T. Richards, T. Shanks, B. J. Boyle, *et al.*, *The 2dF-SDSS LRG and QSO survey : the QSO luminosity function at 0.4 < z < 2.6*, **399**, 1755 (2009), doi:10.1111/j.1365-2966.2009.15398.x, arXiv:0907.2727.

[52] J. Bovy, A. D. Myers, J. F. Hennawi, D. W. Hogg, *et al.*, *Photometric Redshifts and Quasar Probabilities from a Single, Data-driven Generative Model*, **749**, 41 (2012), doi:10.1088/0004-637X/749/1/41, arXiv:1105.3975.

[53] A. J. Ross, W. J. Percival, A. G. Sánchez, L. Samushia, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : analysis of potential systematics*, **424**, 564 (2012), doi:10.1111/j.1365-2966.2012.21235.x, arXiv:1203.6499.

[54] N. Palanque-Delabrouille, C. Yeche, A. D. Myers, P. Petitjean, *et al.*, *Variability selected high-redshift quasars on SDSS Stripe 82*, **530**, A122 (2011), doi:10.1051/0004-6361/201016254, arXiv:1012.2391.

[55] M. R. Blanton, H. Lin, R. H. Lupton, F. M. Maley, *et al.*, *An Efficient Targeting Strategy for Multiobject Spectrograph Surveys : the Sloan Digital Sky Survey "Tiling" Algorithm*, **125**, 2276 (2003), doi:10.1086/344761, astro-ph/0105535.

[56] A. S. Bolton, D. J. Schlegel, É. Aubourg, S. Bailey, *et al.*, *Spectral Classification and Redshift Measurement for the SDSS-III Baryon Oscillation Spectroscopic Survey*, **144**, 144 (2012), doi:10.1088/0004-6256/144/5/144, arXiv:1207.7326.

[57] D. P. Schneider, P. B. Hall, G. T. Richards, M. A. Strauss, *et al.*, *The Sloan Digital Sky Survey Quasar Catalog. IV. Fifth Data Release*, **134**, 102 (2007), doi:10.1086/518474, arXiv:0704.0806.

[58] I. Pâris, P. Petitjean, É. Aubourg, S. Bailey, *et al.*, *The Sloan Digital Sky Survey quasar catalog : ninth data release*, **548**, A66 (2012), doi:10.1051/0004-6361/201220142, arXiv:1210.5166.

[59] I. Pâris, P. Petitjean, N. P. Ross, A. D. Myers, *et al.*, *The Sloan Digital Sky Survey Quasar Catalog : Twelfth data release*, **597**, A79 (2017), doi:10.1051/0004-6361/201527999, arXiv:1608.06483.

[60] P. C. Hewett et V. Wild, *Improved redshifts for SDSS quasar spectra*, **405**, 2302 (2010), doi:10.1111/j.1365-2966.2010.16648.x, arXiv:1003.3017.

[61] Y. Shen, W. N. Brandt, G. T. Richards, K. D. Denney, *et al.*, *The Sloan Digital Sky Survey Reverberation Mapping Project : Velocity Shifts of Quasar Emission Lines*, **831**, 7 (2016), doi:10.3847/0004-637X/831/1/7.

[62] Y. Shen, W. N. Brandt, K. S. Dawson, P. B. Hall, *et al.*, *The Sloan Digital Sky Survey Reverberation Mapping Project : Technical Overview*, **216**, 4 (2015), doi:10.1088/0067-0049/216/1/4, arXiv:1408.5970 [astro-ph.IM].

[63] L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample*, **427**, 3435 (2012), doi:10.1111/j.1365-2966.2012.22066.x, arXiv:1203.6594.

[64] C. P. Ahn, R. Alexandroff, C. Allende Prieto, S. F. Anderson, *et al.*, *The Ninth Data Release of the Sloan Digital Sky Survey : First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey*, **203**, 21 (2012), doi:10.1088/0067-0049/203/2/21, arXiv:1207.7137 [astro-ph.IM].

[65] C. P. Ahn, R. Alexandroff, C. Allende Prieto, F. Anders, *et al.*, *The Tenth Data Release of the Sloan Digital Sky Survey : First Spectroscopic Data from the SDSS-III Apache Point Observatory Galactic Evolution Experiment*, **211**, 17 (2014), doi:10.1088/0067-0049/211/2/17, arXiv:1307.7735 [astro-ph.IM].

[66] L. Anderson, É. Aubourg, S. Bailey, F. Beutler, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples*, **441**, 24 (2014), doi:10.1093/mnras/stu523, arXiv:1312.4877.

[67] B. Reid, S. Ho, N. Padmanabhan, W. J. Percival, *et al.*, *SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12 : galaxy target selection and large-scale structure catalogues*, **455**, 1553 (2016), doi:10.1093/mnras/stv2382, arXiv:1509.06529.

[68] M. E. C. Swanson, M. Tegmark, A. J. S. Hamilton, et J. C. Hill, *Methods for rapidly processing angular masks of next-generation galaxy surveys*, **387**, 1391 (2008), doi:10.1111/j.1365-2966.2008.13296.x, arXiv:0711.4352.

[69] A. J. Ross, S. Ho, A. J. Cuesta, R. Tojeiro, *et al.*, *Ameliorating systematic uncertainties in the angular clustering of galaxies : a study using the SDSS-III*, **417**, 1350 (2011), doi:10.1111/j.1365-2966.2011.19351.x, arXiv:1105.2320.

[70] D. Bianchi et W. J. Percival, *Unbiased clustering estimation in the presence of missing observations*, **472**, 1106 (2017), doi:10.1093/mnras/stx2053, arXiv:1703.02070.

[71] H. A. Feldman, N. Kaiser, et J. A. Peacock, *Power-spectrum analysis of three-dimensional redshift surveys*, **426**, 23 (1994), doi:10.1086/174036, astro-ph/9304022.

[72] E. F. Schlafly et D. P. Finkbeiner, *Measuring Reddening with Sloan Digital Sky Survey Stellar Spectra and Recalibrating SFD*, **737**, 103 (2011), doi:10.1088/0004-637X/737/2/103, arXiv:1012.4804 [astro-ph.GA].

[73] H. Gil-Marín, J. Guy, P. Zarrouk, E. Burtin, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : structure growth rate measurement from the anisotropic quasar power spectrum in the redshift range 0.8 z 2.2*, **477**, 1604 (2018), doi:10.1093/mnras/sty453, arXiv:1801.02689.

[74] W. E. Ballinger, J. A. Peacock, et A. F. Heavens, *Measuring the cosmological constant with redshift surveys*, **282**, 877 (1996), doi:10.1093/mnras/282.3.877, astro-ph/9605017.

[75] S. D. Landy et A. S. Szalay, *Bias and variance of angular correlation functions*, **412**, 64 (1993), doi:10.1086/172900.

Chapitre 3

Theory of large-scale structure formation

Contents

Redshift surveys provide a three-dimensional view of the large-scale structures of the universe whose statistical properties can be studied by modeling of the two-point correlation function. However, since redshift obtained from spectroscopic surveys and from which distances are inferred contain both a contribution from the Hubble expansion and

the LOS velocity, galaxy redshift surveys actually measure a combination of the density and velocity fields in redshift space. Therefore, the two-point correlation in redshift space includes at least three types of non-linearities that are challenging to model theoretically : the non-linear evolution of density and velocity fields, the non-linear relation between dark matter and tracers distribution and the non-linear mapping from real to redshift space.

In this chapter, section 3.1 reviews the dynamics of gravitational instabilities that give rise to the structures seen in galaxy surveys. As long as density fluctuations are small enough, non-linearities can be treated by adopting a perturbative expansion around linear solutions, but this breaks down on small scales and perturbation theories are no longer valid. Section 3.2 thus presents two approaches that have been developed to have insights on the behavior of cosmological fields in the non-linear regime. Both sections are heavily influenced by [1], Will Percival's lectures at the Early Career Scientists eBOSS meeting at the University of Ohio in December 2016 and Sandrine Codis's lectures at the Euclid France summer school in July 2017. Then, section 3.3 deals with the relation between dark matter and tracers statistical properties in the framework of biased tracers. This section is inspired from a series of online lectures given by Franck van den Bosch on the theory of galaxy formation [1] and on talks that have been given at the SnowPAC conference in Utah that I attended in March 2016. Eventually, section 3.4 reviews the two most common methods to model the redshift space distortions and connect the results in real space (theory) and redshift space (observations).

3.1 Development of gravitational instabilities

The standard picture for the formation of large-scale structures that we owe to Lemaitre [2] is that they result from the gravitational amplifications of small primordial density fluctuations. In this scenario, the dynamics of structure formation is mostly driven by the gravitational interaction of collisionless (or at least, weakly-interacting) dark matter particles in an expanding universe (for a textbook, see e.g. [3]). Altough the microscopic nature of dark matter particles has not been identified yet, at scales much smaller than the Hubble radius, all candidates must be cold, hence non-relativistic, before the matter-dominated area so that the damping of small-scale fluctuations due to dark matter is compatible with observed structures [4, 5]. Under these conditions, the equations of motions that describe the evolution of the matter distribution reduce to those of Newtonian gravity. In the WIMP scenario dark matter particles are extremely light compared to the mass scale of galaxies with an expected number density of at least 10^{50} particles per Mpc^3 [e.g. 6]. Therefore, discreteness effects are negligible and collisionless CDM particles can be treated as a fluid that obeys the Vlasov equation for the distribution function in phase space. Moreover, as mentioned in section 1.1.4, we can distinguish two regimes in the formation of structures : a first regime when dark matter particles evolve and collapse into halos, and then the formation of baryonic structures associated with gaseous processes. In this section, we consider the formation of dark matter halos only.

3.1.1 The Vlasov equation

First we assume that the universe is full of dust like particles with same masses m for simplicity and whose only interaction is gravitational. We introduce the phase space density function $f(\mathbf{x}, \mathbf{p}) d^3\mathbf{x} d^3\mathbf{p}$ which represents the number of particles per volume element

1. https://campuspress.yale.edu/astro610/

$d^3\mathbf{x}d^3\mathbf{p}$ where the position $\mathbf{x}$ of the particles is expressed in comoving coordinates and the particle conjugate momentum $\mathbf{p}$ is given by $\mathbf{p} = \mathbf{u}ma$ where a is the expansion factor and u is the peculiar velocity. The peculiar velocity $\mathbf{u}$ is defined by the difference between the total velocity $\mathbf{v}$ and the Hubble flow $\dot{a}\mathbf{x}$ such that :

$$\mathbf{u} = a\frac{d\mathbf{x}}{dt} = \mathbf{v} - \dot{a}\mathbf{x} \tag{3.1}$$

The conservation of particles together with the Liouville theorem that asserts that the volume of the phase-space distribution function is constant with time imply that the total derivative of f vanishes so that :

$$\frac{df}{dt} = \frac{\partial}{\partial t}f(\mathbf{x},\mathbf{p},t) + \frac{d\mathbf{x}}{dt}\frac{\partial}{\partial \mathbf{x}}f(\mathbf{x},\mathbf{p},t) + \frac{d\mathbf{p}}{dt}\frac{\partial}{\partial \mathbf{p}}f(\mathbf{x},\mathbf{p},t) = 0 \tag{3.2}$$

This is the Vlasov equation.

Using the definition of the peculiar velocity given by equation 3.1, we can relate the time variation of the position to the momentum $\mathbf{p}$ by

$$\frac{d\mathbf{x}}{dt} = \frac{\mathbf{p}}{ma^2} \tag{3.3}$$

Then, the time variation of the momentum is generally obtained from the geodesic equation. For distances small compared to the curvature radius of the universe, where the local gravitational potentials are small enough that general relativity effects are negligible, one can treat the growth of perturbations in the local universe in the framework of Newtonian gravity. Therefore, for a particle of velocity $\mathbf{v}$ at position $\mathbf{r}$, the action of all other particles can be treated as a gravitational potential ϕ induced by the local mass density $\rho(\mathbf{r})$

$$\phi(\mathbf{r}) = G\int \frac{\rho(\mathbf{r}'-\mathbf{r})}{|\mathbf{r}'-\mathbf{r}|}d^3\mathbf{r}' \tag{3.4}$$

and the equation of motion yields :

$$\frac{d\mathbf{v}}{dt} = -\nabla_r\phi \tag{3.5}$$

As the peculiar velocity $\mathbf{u}$ represents a perturbation to the background velocity that satisfies the Hubble's law, we can define perturbations to the background values for the density field and gravitational potential :

$$\rho(\mathbf{x},t) = \bar{\rho}(t)(1+\delta(\mathbf{x},t) \tag{3.6}$$

$$\phi(\mathbf{x},t) = \bar{\phi}(\mathbf{x},t) + \Phi(\mathbf{x},t) \tag{3.7}$$

where $\bar{\rho}(t)$ is the spatial average of the density field and the background potential $\bar{\phi}$ satisfies the Poisson equation $\Delta\bar{\phi} = 4\pi G a^2\bar{\rho}$.

So the equation of motion 3.5 becomes :

$$\frac{d\mathbf{v}}{dt} = \frac{d\mathbf{u}}{dt} + \frac{\dot{a}}{a}\mathbf{u} + \ddot{a}\mathbf{x} \tag{3.8}$$

$$= \frac{1}{ma}\frac{d\mathbf{p}}{dt} + \ddot{a}\mathbf{x} \tag{3.9}$$

$$= -\nabla_r\phi = -\frac{1}{a}\nabla_x\phi \tag{3.10}$$

$$= -\frac{1}{a}(\nabla_x\bar{\phi} + \nabla_x\Phi) \tag{3.11}$$

Using the second Friedman equation (equation 1.18) for the homogeneous background and the Gauss theorem ($4\pi G \int \bar{\rho} d^3\mathbf{x} = 4\pi G M$ where M is the mass related to $\bar{\rho}$), one can show that

$$\ddot{a}\mathbf{x} = -\frac{4\pi G}{3} a\bar{\rho} = -\frac{1}{a}\nabla_x \bar{\phi} \tag{3.12}$$

It therefore yields :

$$\frac{d\mathbf{p}}{dt} = -m\nabla_x \Phi(\mathbf{x}, t) \tag{3.13}$$

where Φ is the cosmological gravitational potential which is sourced by the density contrast in the context of metric perturbation in an expanding universe so that :

$$\Delta\Phi(\mathbf{x}) = 4\pi G a^2 \bar{\rho}\delta = \frac{3}{2}\Omega_m(t) H^2(t)\delta \tag{3.14}$$

where the second equality comes from the first Friedmann equation in a flat universe (equation 1.17 with $k = 0$).

Combining equations 3.15, 3.5 and 3.13, we get

$$\frac{\partial}{\partial t} f(\mathbf{x}, \mathbf{p}, t) + \frac{p}{ma^2}\frac{\partial}{\partial \mathbf{x}} f(\mathbf{x}, \mathbf{p}, t) - m\nabla_x \Phi(\mathbf{x}, t)\frac{\partial}{\partial \mathbf{p}} f(\mathbf{x}, \mathbf{p}, t) = 0 \tag{3.15}$$

The system of equations 3.14 and 3.15 form the Vlasov-Poisson equation. This is precisely the set of equations N-body simulations attempt to solve.

The basic conservation equations can now be derived from the first moments of the Vlasov equation. The zeroth-order moment relates the phase-space density to the local mass density field :

$$\rho(\mathbf{x}, t) = \frac{m}{a^3}\int d^3\mathbf{p} f(\mathbf{x}, \mathbf{p}, t) \tag{3.16}$$

The first-order moment defines the mean velocity field (average local velocity of particles in a region of space) u_i and the second-order moment defines the velocity dispersion (also called the stress-tensor) $\sigma_{i,j}$ by :

$$\mathbf{u}_i(\mathbf{x}, t) = \frac{1}{\int d^3\mathbf{p} f(\mathbf{x}, \mathbf{p}, t)}\int d^3\mathbf{p}\frac{p_i}{ma} f(\mathbf{x}, \mathbf{p}, t) \tag{3.17}$$

$$\mathbf{u}_i(\mathbf{x}, t)\mathbf{u}_j(\mathbf{x}, t) + \sigma_{ij}(\mathbf{x}, t) = \frac{1}{\int d^3\mathbf{p} f(\mathbf{x}, \mathbf{p}, t)}\int d^3\mathbf{p}\frac{p_i}{ma}\frac{p_j}{ma} f(\mathbf{x}, \mathbf{p}, t) \tag{3.18}$$

The zeroth moment of the Vlasov equation gives the continuity equation (conservation of mass) by integrating the Vlasov equation w.r.t $\mathbf{p}$

$$\frac{\partial\delta(\mathbf{x}, t)}{dt} + \frac{1}{a}\nabla\cdot[(1 + \delta(\mathbf{x}, t))\mathbf{u}(\mathbf{x}, t)] = 0 \tag{3.19}$$

The next two moments gives another equation of conservation : integrating the Vlasov equation w.r.t. $\mathbf{p}$ after multipling by $\mathbf{p}$, then subtracting $\bar{\rho}\mathbf{u}(\mathbf{x}, t)$ times the continuity equation (equation 3.19) gives the Euler equation which describes the conservation of momentum :

$$\frac{\partial\mathbf{u}_i(\mathbf{x}, t)}{\partial t} + \frac{\dot{a}}{a}\mathbf{u}_i(\mathbf{x}, t) + \frac{1}{a}\mathbf{u}_j(\mathbf{x}, t)\cdot\nabla_j\mathbf{u}_i(\mathbf{x}, t) = -\frac{1}{a}\nabla_i\Phi(\mathbf{x}, t) - \frac{1}{\rho a}\nabla_j\rho\sigma_{ij}(\mathbf{x}, t) \tag{3.20}$$

It is worth noting that this equation is very similar to the one in hydrodynamics for perfect fluids with in our case two additional terms, one which accounts for the expansion of the

universe and another one for the anisotropic stress tensor related to the pressure of the fluid.

At the early stages of the development of gravitational instabilities, the velocity dispersion due to peculiar motions is much smaller than the velocity gradients induced by the density fluctuations at the scales of interest. One can thus assume that the cosmological structure formation is driven by matter with negligle velocity dispersion and pressure (single-flow approximation) such that $\sigma_{ij} = 0$ and that all particles have identical peculiar velocities. This assumption is valid at the beginning of structure formation when structures had no time to collapse and virialize. This is why N-body simulations start at $z \sim 100$, when the single-flow approximation and the linear regime are still a satisfactory description of the dark matter dynamics. Later stages involve the superposition of several streams which will break down this approximation and lead to the generation of velocity dispersion due to multiple streams, which is called *shell-crossing*.

3.1.2 Growth of structure in linear theory

Using the single-flow regime of the Vlasov-Poisson equation that describes the development of gravitational instabilities in a pressureless fluid at the beginning of structure formation, one can linearize the equations of conservation by assuming that the amplitude of the density fluctuations are small and that the velocity gradients are small compared to the Hubble constant :

$$\delta((\mathbf{x}, t) \ll 1, \tag{3.21}$$

$$\frac{1}{a}\nabla_j \mathbf{u}_i(\mathbf{x}, t) \ll H(t) \tag{3.22}$$

This approximation corresponds to the linear regime where the equations of conservation become

$$\frac{\partial \delta(\mathbf{x}, t)}{\partial t} + H\theta(\mathbf{x}, t) = 0, \tag{3.23}$$

$$\frac{\partial \theta(\mathbf{x}, t)}{\partial t} + H(t)\theta(\mathbf{x}, t) + \frac{3}{2}\Omega_m(t)H^2(t)\delta(\mathbf{x}, t) = 0 \tag{3.24}$$

where $\nabla\Phi$ has been replaced using the Poisson equation (equation 3.14) and $\theta = \frac{1}{aH}\nabla_x \cdot \mathbf{u}(\mathbf{x}, t)$ is the divergence of the velocity field. It completely describes the velocity field since the vorticity field is negligible in the single-flow approximation and any initial vorticity rapidly decays due to the expansion of the universe. In the non-linear regime, where the single-flow approximation is no longer valid, [7] showed that the presence of velocity dispersion can generate vorticity.

Then taking the divergence of equation 3.24 and replacing the time derivative of the density contrast by its expression in equation 3.23, we obtain the evolution of the density contrast that describes the growth of perturbations in the linear regime :

$$\frac{\partial^2 \delta(\mathbf{x}, t)}{\partial t^2} + H(t)\frac{\partial \delta(\mathbf{x}, t)}{\partial t} = 4\pi G a^2 \bar{\rho}\delta(\mathbf{x}, t) = \frac{3}{2}\Omega_m(t)\delta(\mathbf{x}, t) \tag{3.25}$$

This second order dynamical equation is linear and does not present any dependency on the spatial coordinates. Therefore, in the linear regime for a pressureless fluid, the linear growth rate of the fluctuations is scale-independent. It allows us to look for solutions by decoupling the spatial and time contribution and since it is a second-order differential equation, it has two independent solutions, i.e. $\delta(\mathbf{x}, t) = D^{(+)}(t)\delta_+(\mathbf{x}, 0) + D^{(-)}(t)\delta_-(\mathbf{x}, 0)$.

The spatial functions only depend on the initial density field configuration while the time-dependency is related to the energy content of the universe. There is no unique solution for all the cosmological models. If we consider the late-time universe with a mix of matter, cosmological constant and a curvature term, the first Friedman equation gives for the Hubble parameter :

$$H = H_0\sqrt{\Omega_{m,0}a^{-3} + (1 - \Omega_{m,0} - \Omega_{\Lambda,0})a^{-2} + \Omega_{\Lambda,0}} \tag{3.26}$$

It can be checked that this expression is a solution of equation 3.25. It corresponds to the decaying mode, $D^{(-)}(t) \propto H(t)$, and the other solution is found to be $D^{(+)}(t) \propto H(t)\int \frac{dt}{(aH)^2}$ [1]. The $D^{(+)}(t)$ solution is the growing mode which leads to structure formation, it is called the **growth function**. In what follows, we consider the growing solution only.
Moreover, the continuity equation allows us to find an explicit relation between the density contrast and the divergence of the velocity field :

$$a\frac{\partial\delta(\mathbf{x},t)}{\partial a} + \theta(\mathbf{x},t) = 0 \tag{3.27}$$

Therefore, we can write θ so that

$$\theta(\mathbf{x},t) = \frac{\partial \ln D^+}{\partial \ln a}\delta_+(\mathbf{x},t) \tag{3.28}$$

We call f, the **linear growth rate of structure**, the factor of proportionality between θ and δ and it thus corresponds to the logarithmic derivative of the growing mode of the density contrast :

$$f = \frac{d \ln D^{(+)}}{d \ln a} \tag{3.29}$$

So the linear regime is characterized by a linear coupling the matter velocity and density fields so that : $\theta_m = f\delta_m$.

For an Einstein-de Sitter universe ($\Omega_m = 1$ and $\Omega_\Lambda = 0$), $f = 1$.
For a flat universe, $\Omega_m + \Omega_\Lambda = 1$, with a cosmological constant, a good parametrization [8] is $f(\Omega_m) = \Omega_m^\gamma$ where the growth index γ is related to the equation of state of dark energy by [9, 10] :

$$\gamma = \frac{3(1 - w_{\rm DE})}{5 - 6w_{\rm DE}} \tag{3.30}$$

This is where the index $\gamma = 0.55$ comes from for a flat Λ-CDM cosmological model based on general relativity. Therefore, equation 3.30 shows how growth rate measurements can constrain the nature of dark energy.
For a universe without cosmological constant, a good parametrization is $f(\Omega_m) = \Omega_m^{\frac{3}{5}}$.Both parametrizations as a function of Ω_m are shown in figure 3.1.

3.1.3 Lagrangian point of view

So far, the previous approach dealt with density and velocity fields as function of spatial coordinates wihch describe individual over-densities that stay at a fixed position and whose amplitude grow or decay. This is the Eulerian point of view. But it is possible to work in another framework, the Lagrangian scheme, where we follow the trajectories of particles or fluid elements. In this framework, the fields are expressed as functions of the initial particle positions instead of fixed comoving coordinates. In the Lagrangian

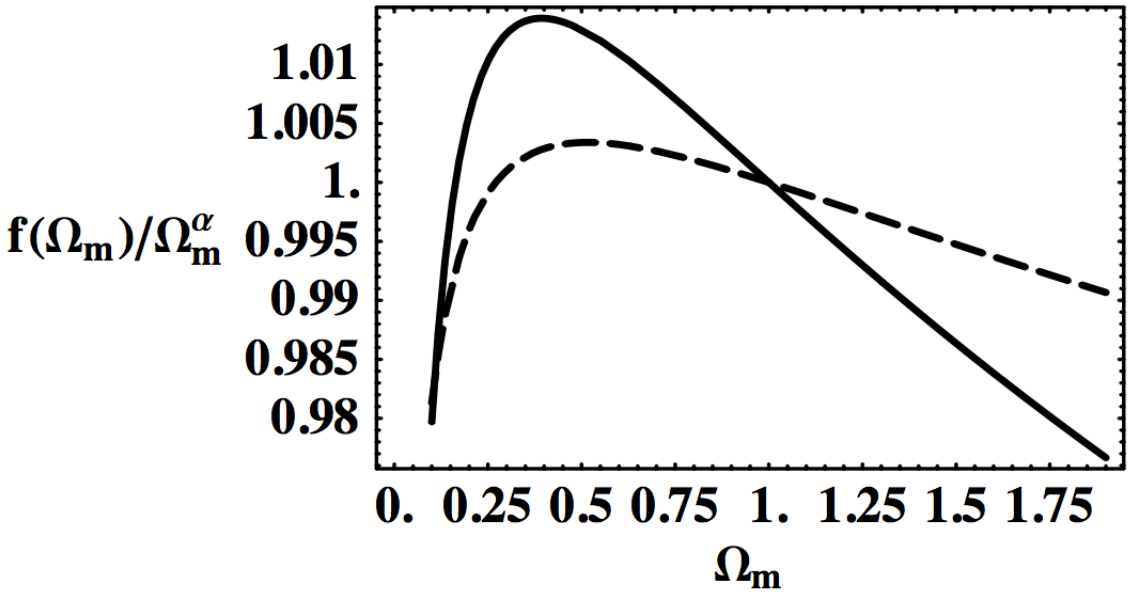


FIGURE 3.1 – Growth of structure as defined by equation 3.29 divided by its estimated value for a flat universe with Λ (dashed line) and for a universe without Λ (solid line) as a function of Ω_m. Figure from [1]

framework, the object of interest is the displacement field $\Psi(\mathbf{q}, t)$ which maps the initial particle position $\mathbf{q}$ to the final comoving (Eulerian) particle position $\mathbf{x}$:

$$\mathbf{x}(\mathbf{q}, t) = \mathbf{q} + \mathbf{\Psi}(\mathbf{q}, t) \ . \tag{3.31}$$

The continuity equation that states the conservation of mass therefore gives :

$$\rho(\mathbf{x}, t)d^3\mathbf{x} = \rho(\mathbf{q})d^3\mathbf{q} \tag{3.32}$$

The equation of motion becomes :

$$\mathbf{u} = a\dot{\mathbf{x}} = a\dot{\Psi} \tag{3.33}$$

$$\dot{\mathbf{u}} + H\mathbf{u} = -\frac{1}{a}\nabla_x\Phi \tag{3.34}$$

where $\dot{\mathbf{u}}$ is the time derivative when $\mathbf{q}$ is fixed, Φ is the cosmological gravitational potential introduced in equation 3.13 and ∇_x is the gradient operator in Eulerian comoving coodinates $\mathbf{x}$. Using equation 3.33, one can rewrite equation 3.34 for ψ :

$$\ddot{\Psi} + H\dot{\Psi} = -\frac{1}{a^2}\nabla_x\Phi \tag{3.35}$$

Taking the divergence of equation 3.34 w.r.t $\mathbf{x}$ and using the Poisson equation, one finally gets the equation that describes the evolution of the displacement field Ψ :

$$\nabla_x\ddot{\Psi} + H\nabla_x\dot{\Psi} = 4\pi G(\rho - \bar{\rho}) \tag{3.36}$$

The major difficulty in the Lagrangian description is that the non-linearities are encoded in the mapping from Eulerian to Lagrangian coordinates and in the relation between the

displacement field and the local density field. It therefore implies to manipulate operators w.r.t Eulerian coordinates $\mathbf{x}$ whose dependence with Lagrangian coordinates $\mathbf{q}$ is expressed in equation 3.31.

However, in the linear regime, [11] showed that one can consider the linear solution for the displacement field ψ while keeping the continuity equation given by equation 3.32. In this approximation, called the Zel'dovich approximation, the divergence operator ∇_x can be assimilated to ∇_q at zeroth order. Equation 3.36 becomes similar to the linearized equation 3.25 that describes the growth of perturbations in the linear regime.
The equation for Ψ has also two independent solutions, i.e $\Psi(\mathbf{q},t) = D^{(+)}(t)\Psi_+(\mathbf{q}) + D^{(-)}(t)\Psi_-(\mathbf{q})$. The **Zel'dovich solution** thus corresponds to the growing mode : $\Psi(\mathbf{q},t) = D^{(+)}(t)\Psi_+(\mathbf{q})$ and it is given by :

$$\Psi(\mathbf{q}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{q}} \frac{i\mathbf{k}}{k^2}\delta_l(k) \tag{3.37}$$

[11] showed that for the Jacobian between Eulerian and Lagrangian coordinates can be expressed as :

$$J = |\frac{d\mathbf{x}}{d\mathbf{q}}| = |\delta^K_{ij} + \frac{\partial\Psi_i}{\partial q_j}| \tag{3.38}$$

And if the initial density distribution is homogeneous, we have $\rho = \bar{\rho}/J$, so using the growing solution for ψ, we obtain for the density :

$$\rho = \frac{\bar{\rho}}{[1 - D^{(+)}(t)\lambda_1][1 - D^{(+)}(t)\lambda_2][1 - D^{(+)}(t)\lambda_3]} \tag{3.39}$$

where λ_i are the eigenvalues of the deformation tensor $\partial\psi_i/\partial q_j$. If $\lambda_i > 0$, it implies a collapse in the direction of the i^{th} eigenvector, if $\lambda_i < 0$ it implies expansion. The Zel'dovich approximation thus allows us to understand the collapse of structures by describing the anisotropic collapse of structures that give rise to the cosmic web : it starts from walls ($\lambda_1 > 0$), then filaments ($\lambda_{1,2} > 0$) and eventually knots $\lambda_{1,2,3} > 0$.
As long as $D^{(+)}(t)\lambda_i \ll 1$, the growth of perturbations is well described by the linear theory and when $D^{(+)}(t) = 1/\lambda_i$, shell crosssing becomes dominant along the direction of the i^{th} eigenvector. the Zel'dovich solution is no longer a good description.

3.1.4 Beyond linear theory

As mentioned above, the linear regime is valid as long as $\delta((\mathbf{x},t) \ll 1$, or in other words, the variance of the linear fluctuations is treated as a small parameter close to unity. The domain of validity of the linear regime thus corresponds to scales above $50-60h^{-1}$Mpc. For scales below, non-linearities arise because of non-linear couplings in equations 3.19, 3.20 and 3.14 when considering all the terms without linearization. The presence of non-linear couplings is particularly clear in Fourier space after rewriting these equations, using the convention given by equation 1.40. Indeed, at large scales, in the linear regime when fluctuations are small, one can see that different Fourier modes evolve independently so that it is more natural to work in Fourier space. Then when non-linear terms in the perturbation series are taken into account, couplings between different Fourier modes appear.

The continuity equation thus becomes :

$$\frac{\partial\delta(\mathbf{k},t)}{\partial t} + H\theta(\mathbf{k},t) = \int d^3\mathbf{k}_1 d^3\mathbf{k}_2 \delta_{\rm D}(\mathbf{k}-\mathbf{k}_{12})\alpha(\mathbf{k}_1,\mathbf{k}_2)\theta(\mathbf{k}_1,t)\delta(\mathbf{k}_2,t) \tag{3.40}$$

Then taking the divergence of the Euler equation and using the Poisson equation to replace $\nabla^2\Phi$, we get in Fourier space :

$$\frac{\partial\theta(\mathbf{k},t)}{\partial t}+H\theta(\mathbf{k},t)+\frac{3}{2}\Omega_m H^2(t)\delta(\mathbf{k},t)=-\int d^3\mathbf{k}_1 d^3\mathbf{k}_2\delta_{\rm D}(\mathbf{k}-\mathbf{k}_{12})\beta(\mathbf{k}_1,\mathbf{k}_2)\theta(\mathbf{k}_1,t)\theta(\mathbf{k}_2,t) \tag{3.41}$$

where α and β are defined by :

$$\alpha(\mathbf{k}_1,\mathbf{k}_2)=\frac{\mathbf{k}_{12}\cdot\mathbf{k}_1}{k_1^2}\text{ and }\beta(\mathbf{k}_1,\mathbf{k}_2)=\frac{\mathbf{k}_{12}(\mathbf{k}_1\cdot\mathbf{k}_2)}{2k_1^2k_2^2} \tag{3.42}$$

where $\mathbf{k}_{12}=\mathbf{k}_1+\mathbf{k}_2$. These functions are called the fundamental mode coupling functions and encode the non-linear couplings between different Fourier modes that lead to a non-linear evolution of the cosmological fields.

In the "weakly" non-linear regime, also called the quasi-linear regime (between 20 and 60 h^{-1}Mpc), the ability of perturbation theory models to interpret results from large-scale structure surveys is based on the fact density fluctuations become small enough at these intermediate scales that a perturbative approach suffices to understand their evolution.

Eulerian perturbation theory In the Eulerian framework, it is therefore possible to expand the density and velocity fields about the linear solution :

$$\delta(\mathbf{x},t)=\sum\delta^{(n)}(\mathbf{x},t) \tag{3.43}$$

$$\theta(\mathbf{x},t)=\sum\theta^{(n)}(\mathbf{x},t) \tag{3.44}$$

where $\delta^{(1)}$ and $\theta^{(1)}$ are linear in the initial density field, more precisely $\delta^{(1)}=$ as explained above, then $\delta^{(2)}$ and $\theta^{(2)}$ are quadratic in the initial field, etc. The n$^{\rm th}$-order contributions to δ (resp. θ) can thus be expressed as functions of the linear solution and of functions called "kernels" which are constructed from the fundamental mode coupling function α (resp. β). Non-linear couplings thus imply higher-order corrections to the linear theory prediction. Using the diagrammatic representation, those corrections correspond to "loops" that are added to the configuration given by the linear solution which is called the tree-level.

The formalism beyond the linear perturbation theory involves theoretical objects such as kernels that encode the modes coupling, but also a diagramatic representation with propagators and vertices as developed in quantum field theory which are beyond the scope of this thesis. We just provide a diagrammatic representation of the n$^{\rm th}$-order contribution to $\delta(k)$ in figure 3.2 where the open circles denote factors of δ_0 and the vertex denotes a momentum-conserving integral of F_n over intermediate wavevectors q_i. The F_n (resp. G_n) for δ (resp. θ) function is constructed from the fundamental mode coupling function $\alpha(\mathbf{k}_1,\mathbf{k}_2)$ (resp. $\beta(\mathbf{k}_1,\mathbf{k}_2)$) and encodes the non-linear couplings. Moreover, a variety of methods have been developed to model the galaxy clustering statistics on intermediate quasi-linear scales (20-80 h^{-1}Mpc) in EPT (for a review, see e.g. [1]) that includes standard PT, renormalized PT [12, 13, 14, RPT] and regularized PT [15, RegPT] for the most used ones.

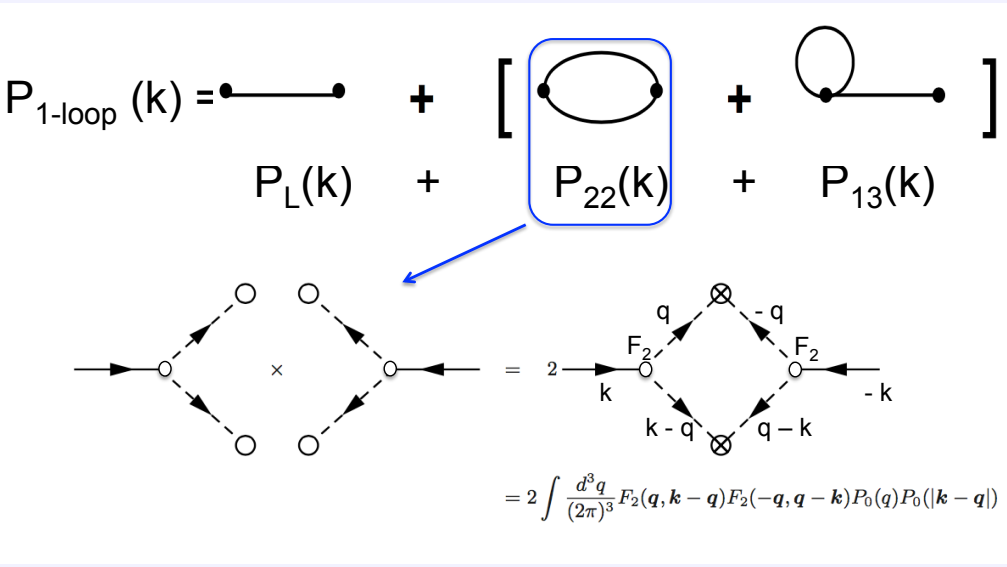
1-loop correction to the matter power spectrum

The first correction to the linear matter power spectrum is obtained by summing over all possible pairing of open circles up to δ_2 where open circles are paired according to the rule

This diagrammatic representation corresponds to the linear matter power spectrum, dubbed the tree-level.
The first correction is second-order in the initial power spectrum, fourth order in the initial density contrast and 1-loop in the diagrammatic representation. It is given by

$$P_{1-\text{loop}}(k) = P_L(k) + P_{22}(k) + P_{13}(k) \tag{3.45}$$

whose diagrammatic representation includes a diagram with one loop



The contribution $P_{22}(k)$ is detailed where the vertices correspond to F_2 and the overall factor 2 comes from the two equivalent ways of pairing the open circles. After invoking momentum conservation at vertices and translational invariance of the 2-point statistics, only a single wavevector remains to be integrated. In general, all diagrams contributing to $P^{(n)}$ contain $n-1$ loops, requiring integration over $n-1$ independent wavevectors.

Lagrangian perturbation theory Alternatively, In the Lagrangian prescription, as the object of interest is the displacement field ψ, the perturbative expansion can be written as :

$$\psi(\mathbf{q}, t) = \sum \psi^{(n)}(\mathbf{q}, t) \tag{3.46}$$

where $\psi^{(1)}$ corresponds to the Zel'dovich solution given by equation 3.37 which is the linear order. [17] proposed a new resummation of the non-linear terms that extended the domain of validity compared to LPT, in what follows we will refer to this technique as LRT (Lagrangian Resummation Theory).

[16, 18] compared different perturbation theory codes, i.e. linear PT, LRT and 1-loop

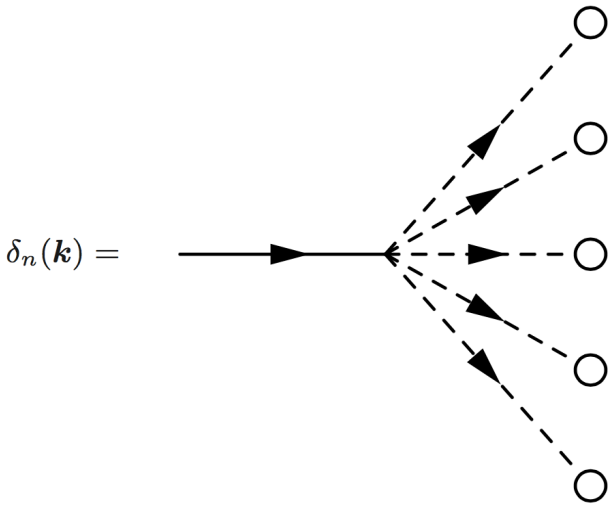


FIGURE 3.2 – Diagrammatic representation of the n$^{\text{th}}$ contribution to $\delta(k)$. From [16].

/ 2-loop SPT and RPT, at the level of halos in real space using N-body simulations. They found similar conclusions :

— Linear PT is no longer valid at $k \geq 0.10\,h\text{Mpc}^{-1}$ for all redshifts.

— In general, working at 2-loop correction improves the accuracy of the predictions without a very high price in terms of computational ressources

— For 2-loop SPT and RPT and for LRT, 1% accuracy can be achieved on the multipoles of the matter power spectrum with $k_{\text{max}} \sim 0.10\,h\text{Mpc}^{-1}$ at $z = 0$ and $k_{\text{max}} \sim 0.20\,h\text{Mpc}^{-1}$ at $z = 1$.

In section 3.4, we will see that these PT predictions have to be obtained in redshift space to be compared with observations. It thus implies additional non-linearities that come from the non-linear mapping between real and redshift space.

3.2 Non-linear clustering of dark matter

The simplest way to go beyond linear theory is the brute-force simulation of gravitational structure formation using Newtonian dynamics of a system of particles. Such N-body simulations solve the Vlasov-Poisson equations given in equations 3.14 and 3.15 by partitioning the cold dark matter density field into particles, each corresponding to a typical mass of $m_p \simeq 10^{10} M_\odot$. The numerical code is usually heavily parallelized and uses adaptive time discretization to solve the dynamical equations. Section 3.2.1 provides a description of numerical simulations and the ones we use for the analysis of the eBOSS quasar sample. But the simulation of the non-linear evolution with the N-body technique is computationally very expensive, so parallel analytical approaches have been developed to model some statistical properties beyond the linear regime. One of the most common method is the **halo model** [19, 20, 21] which is an analytical approach to describe the non-linear clustering by combining the halo mass function with a halo density profile to obtain predictions for the clustering statistics. The ingredients of the halo model are described in section 3.2.2.

3.2.1 Numerical simulations

Cosmological dark matter simulations have been developed to understand the formation of structures (for further details, see [22, 23]), and in particular to explore the non-linear regime of the gravitational evolution of structures in the universe where perturbation theories break down (typically it happens on scales smaller than a few Mpc). In such fictive realizations of the universe, the density field is represented by the sum of a set of discrete particles. Current state-of-the-art numerical simulations can follow the dynamics of about 10^{10-12} particles, which although impressive, is still tens orders of magnitude smaller than the number of dark matter particles expected in a cosmological volume.

3.2.1.1 Production of N-body simulations

Instead of solving the Vlasov equation for collisionless dark matter which would imply a 6D grid with too many cells, sampling techniques are used to divide the volume into N elementary volumes, and the density field can be considered as the sum of N particles (or 'bodies') with positions, velocities and same mass.
The initial conditions for N -body simulations are generated randomly according to the linear theory power spectrum (as obtained with CAMB for instance and given by equation 1.50) for a moderately high redshift. Usually this starting point is chosen to be around $z \sim 100$, because if the perturbations are too small, the N -body technique is inefficient, but for too late times, the perturbation theory model becomes invalid.
The basic steps in an N-body simulation can be summarized as follows [24] :

(i) implementation of initial conditions at high enough time redshift so that $\Omega_m = 1$ (matter-dominated era)

(ii) calculation of the force by solving the Poisson equation

(iii) update the positions and velocities of particles over time steps

(iv) cross-checks (e.g. tests of energy conservation)

(v) go back to (ii) until simulation is completed

The equations of motion for collisionless N-body particles can be written :

$$\frac{d\mathbf{v_i}}{dt} = \sum_{i \neq j}^{N} G m_j \frac{\mathbf{r}_j - \mathbf{r}_i}{(|\mathbf{r}_j - \mathbf{r}_i|^2 + \epsilon^2)^{3/2}} \tag{3.47}$$

where ϵ is called the softening parameter. For $\epsilon = 0$, the equations reduce to the standard Newtonian equations of motion. The main for setting $\epsilon \neq 0$ is to suppress the $1/r$ singularity in the Newtonian potential, which simplifies the numerical integrations of these equations a lot [e.g 25]. The simplest way of calculating the force consists in summing over all other particles. Such method is called particle-particle (PP or direct summation) and provides robust and accurate results. But this direct summation is very CPU consuming where the cost of computing forces on all particules requires $O(N^2)$ operations. Therefore, there is a compromise to be found between accuracy and speed and the characteristics of the N-body simulation will depend on the implementation of the force solver and time stepper. In what follows, I briefly present the most important and commonly used discretization techniques :

— tree algorithm : it consists in decomposing hierarchically the system on a tree structure that is made-up of several clusters of particles. Therefore, the list of interactions that the clusters act on each particle is much shorter than when considering PP method, resulting in a $O(N \log N)$ code instead of $O(N^2)$.

— particle-mesh (PM) algorithm : the mass of each particle is interpolated on a fixed grid of N_c^3 cells to compute the density. The Poisson equation is solved on a grid generally using a fast Fourier transform (FFT) taking advantage of the periodic boundary conditions, then forces are interpolated back on the particles. It requires less memory than the tree code but its spatial resolution which is fixed by the size of the grid N_c is lower.

— hybrid methods have been developed to increase the spatial resolution of the PM technique such as by adding a short-range contribution obtained by direct summation of individual interactions between nearby particles (PP + PM code) or by using an adaptive mesh refinement where the mesh can be locally increased when required thanks to a hierarchy of sub-grids.

A complete description of N-body techniques to solve the Vlasov-Poisson equation is beyond the scope of this thesis. For a more detailed overview, see e.g. [23, 24] and the associated articles when referring to a specific N-body simulation.

3.2.1.2 Halo finder

Snapshots of the evolved density field are stored to disk and usually post-processed to extract information about bound structures, the DM halos. These halos are gravitationally bound systems of DM particles, where the relation between kinetic and potential energy was driven to an equilibrium state by gravitational virialization. More information on halos formation and properties will be provided in section 3.2.2.
Here we just described the most common algorithms used to define halos once the dark matter particles have been formed within the simulation :

— Bound Density Maxima [BDM, 26, 27] : The algorithm detects local density maxima, determines a spherical cut-off for the halo and removes unbound particles from the halo, i.e. those particles with a speed that exceeds the escape velocity. The latest version even allows to detect if a halo is a subhalo (i.e. its center lies within the virial radius of a larger halo) or a distinct halo. Please note, that distinct halos may still overlap and subhalos are not necessarily lying completely within their host, only their centre.

— Friends-of-friends [FOF, 28, 29] : Particles are "linked" together if their distance lies below a certain threshold, called "linking length". This means that the distances of particles at the boundary of such a linked object (a "FOF group") are smaller than or equal to the linking length, corresponding to a density threshold.

— Robust Overdensity Calculation using K-Space Topologically Adaptive Refinement [Rockstar, 30] is based on an adaptative hierarchical refinement of FoF groups to take into 3D-velocity components and time, in addition to 3D-position components. It thus uses a 6D-linking length to define FoF groups which makes the identification of dark matter halos and substructures much more robust. It also allows us to reconstruct the history of a halo and its inheritance thanks to the time-component.

The left panel of figure 3.3 is an illustration of the BDM halo finder where the centers of distinct halos do not lie within the virial radius of a larger halo, in contrast to subhalos, which allows the algorithm to differenciate them. The right panel shows FOF groups with different linking lengths where substructures correspond to FOF groups with smaller linking lengths that lie inside a larger host halo.

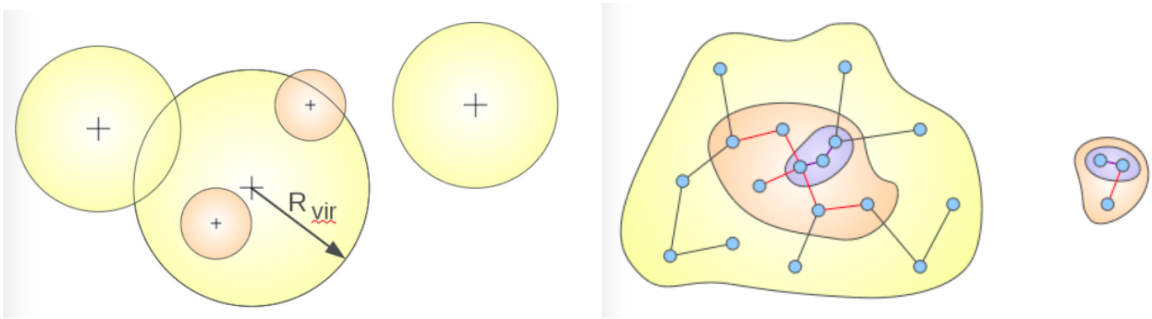


FIGURE 3.3 – Illustration of the BDM (left panel) and the FOF (right panel) halo finder algorithms. Figures from the CosmoSim database.

N-body simulations for the DR14 eBOSS quasar sample

Previous analyses of the quasar clustering have studied the properties of dark matter halos that host quasars. Such studies have been led using the quasar sample of SDSS-III BOSS [31, 32] and a preliminary sample of SDSS-IV eBOSS quasars [33, 34], they confirmed the picture that quasars in the eBOSS redshift range reside in dark matter halos of mass $M \sim 10^{12.5} M_{\odot}$.
The typical halo mass and the maximum separation used for clustering analysis set the requirements in terms of typical for dark matter particles and box size a N-body simulation needs to fulfull. In the analysis of the DR14 quasar sample, we use different sets of N-body simulations to test the domain of validity of the perturbation theory predictions :

— **MultiDark simulations** The MultiDark simulation MDPL2 and BigMDPL [35] were used at an earlier stage to get familiar with the ingredients of the RSD model. They can be found in the CosmoSim database [a]. These simulations have been run with the code GADGET-2 [36] which is a solver based on the tree method for short-range gravitational forces while long-range forces are computed with a FFT-based PM scheme. It was runned in a flat ΛCDM model with $\Omega_m = 0.307$, $\Omega_l = 0.693$, $\Omega_b = 0.048$, $n_s = 0.96$, $h = 0.677$ and $\sigma_8 = 0.8228$. The initial conditions, based on initial Gaussian fluctuations, for the simulation are generated with the Zel'dovich approximation at $z_i = 100$. More information on the volume, number of particles and mass resolution are provided in table 3.1. Given the typical range of masses of halos hosting quasars, the mass of dark matter particles in the BigMDPL simulation is too big to have well-resolved halos (for instance, halos of mass below 10^{12} contain less than 100 particles). The MDPL simulation has the required mass resolution but the box size is too small to be precise enough at the BAO scale ($\sim 100\,h^{-1}$Mpc). These simulations are thus at the limit for being used for the eBOSS quasar sample [37].

— **Outer Rim simulation** The Outer Rim simulation has been carried out with HACC (Hardware/Hybrid Accelerated Cosmology Code) presented in [38]. Information on the volume, number of particles and mass resolution of the Outer Rim simulation are provided in table 3.1. The cosmological parameters are : $\Omega_{\rm cdm} h^2 = 0.1109$, $h = 0.71$, $\sigma_8 = 0.8$, $\Omega_{\rm b} h^2 = 0.02258$, and $n_s = 0.963$, and are consistent with the WMAP7 cosmology [39]. The initial conditions are calculated at $z = 200$ using the Zel'dovich approximation.

a. https://www.cosmosim.org

TABLE 3.1 – Main characteristics of the N-body simulations used for the analysis of the eBOSS quasar sample : mass resolution in $[M_\odot h^{-1}]$, box size in $[h^{-1}\mathrm{Mpc}]$

simulation	cosmology	solver	# particles/Mpc^3	mass resolution	box size	halo finder
MDPL	Planck	Gadget-2	$(3840)^3$	1.51×10^9	1000	Rockstar
BigMDPL	Planck	Gadget-2	$(3840)^3$	2.36×10^{10}	2500	Rockstar
Outer Rim	WMAP7	HAAC	$(10240)^3$	1.9×10^9	3000	FOF

3.2.2 The halo model

Simulations have shown that an initially smooth matter distribution evolves into a complex network of sheets, filaments and knots. The dense knots are often called dark matter halos. They correspond to the first structures of matter that have undergone a gravitational collapse. In the collisionless dark matter assumption, one can describe the complex distribution of dark matter with halos whose mass function is derived from simulations and with a profile for dark matter within halos. These ingredients are also the basics of the halo model whose aims at providing an analytical framework to describe the formation and evolution of halos, these highly non-linear objects, which are traditionally studied using numerical simulations. The approach assumes that all the mass in the universe is distributed into distinct dark matter halos, which can be seen as basic units.

3.2.2.1 Spherical top-hat collapse

Although the equations of the dynamics are highly non-linear, the spherical collapse model provides a simple explanation for the formation of dark matter halos assuming they have been formed from an initial spherical perturbation. Its application to an initially top-hat density perturbation was first studied by [40] and then by [3, 41]. We consider the universe as homogeneous, except for a single top-hat spherical perturbation that is embedded in the background universe, in its matter-dominated era so that we work within Einstein-de Sitter cosmology ($\Omega_m = 1$, $\Omega_\Lambda = 0$, $k = 0$).

Let R_i and δ_i denote the radius and the over-density of the initial perturbation at t_i and let $\bar{R}_i$ and $\bar{\rho}_i$ denote the radius and density of the background at the same time. Because of mass conservation, the mass within the spherical perturbation is :

$$M = \frac{4}{3}\pi R_i^3 \bar{\rho}_i[1+\delta_i] = \frac{4}{3}\pi R_t^3 \bar{\rho}_t[1+\delta_t] \tag{3.48}$$

The Birkhoff's theorem (also called the Newtonian theorem in the non-relativistic limit) states that a spherically matter distribution outside a sphere exerts no force on that sphere and vice-versa the background expands without being disturbed by the spherical perturbation. Under these assumptions, the background and the perturbation are entirely decoupled and the application of the Birkhoff's theorem to the isolated spherical perturbation yields the following equation of motion :

$$\frac{d^2R}{dt^2} = -\frac{GM}{R^2} \tag{3.49}$$

The conservation of energy that one can obtain after integrating the equation of motion once yields :

$$\frac{1}{2}\left(\frac{dR}{dt}\right)^2 - \frac{GM}{R} = E \tag{3.50}$$

where E is the total energy of the spherical perturbation. The gravitationally bound situation which implies a collapse of the spherical perturbation corresponds to $E < 0$. In

this case, the solution takes the following form :

$$R = R_\star(1 - \cos\tau) \tag{3.51}$$
$$t = t_\star(\theta - \sin\tau) \tag{3.52}$$

with $R_\star = \frac{GM}{2|E|}$ at the time $t_\star = \frac{GM}{(2|E|)^{3/2}}$. This solution implies the following evolution of the spherical perturbation that is illustrated in figure 3.4 :

(i) linear growth in accordance with the expansion of the background universe where δ_{lin} corresponds to the prediction for the linear density $\delta_{\rm lin} \propto D(a) \propto a \propto t^{2/3}$.

(ii) turnaround at $\tau = \pi$: as time passes, the perturbation grows up to its maximal size $R_{max} = R(\tau = \pi) = 2R_\star$ and then stops expanding. At this stage, the perturbation grows to an over-density $\Delta_{\rm turnaround} = \delta_{\rm turnaround} + 1 \simeq 5.55$ and the linear density grows to $\delta_{\rm turnaround} = 1.06$.

(iii) virialization at $\tau = 2\pi$: it corresponds to a dissipative effect that converts the kinetic energy of collapse into random motions and finally gives rise to a virialized object of a given size that we call *halo*. If the perturbation were perfectly symmetric, the over-density would actually collapse to a single point becoming infinitely dense. Instead, shell crossing occurs leading to exchanges of momentum angular. Thus, the system relaxes towards a virial equilibrium which is reached when $2E_{\rm kinetic} + E_{\rm potential} = 0$ (virial theorem). It implies that the perturbation collapses up to a radius of about half of its maximal size, which corresponds to its virial radius $R_{\rm vir} = R_{\rm max}/2$. At this stage, the perturbation grows to an over-density $\Delta_{\rm vir} = \delta_{\rm vir} + 1 \simeq 178$ and the linear density grows to $\delta_{\rm collapse} = 1.69$.

Thus, this simple model provides useful insights in the non-linear evolution of spherical perturbations. In particular, we have shown that perturbations form gravitationally-bound dark matter halos when density perturbation become 150-200 times denser than the background (which corresponds to $\Delta_{\rm vir} \simeq 178$). The exact value of the critical density $\delta_c = \delta_{\rm collapse}$ required for spherical collapse depends on cosmology, for a universe with a cosmological constant for instance, it is larger because the late-time mean cosmological density is lower than expected (due to more expansion). Although the above treatment is only valid for an EdS cosmology, similar models can be constructed for other cosmologies including ΛCDM and very similar results for the density at virialization can be found [42].

3.2.2.2 Halo density profile

Simulations show that the end state of virialization is a halo with a centrally-concentrated mass distribution. But if massive halos correspond to higher peaks in the initial density field, the density run around a high peak is shallower than the run around a smaller peak [43], so high peaks are less centrally-concentrated and we can expect that massive virialized halos have a less centrally-concentrated mass distribution compared to low-mass halos. This is what simulations tend to show, with a density profile well-fitted by a double power-law function. Several power-law density profiles have thus been proposed, one of the most frequently used in cosmology are the Navarro-Frenk-White (NFW, [44] and the Hernquist [45] profiles. Let $\rho(R|M)$ denote the halo density profile :

$$\rho(R|M) = \frac{\rho_s}{(R/R_s)^\alpha(1 + R/R_s)^\beta} \tag{3.53}$$

where R_s is the core radius and ρ_s the density at that radius. Setting $(\alpha, \beta) = (1, 3)$ and $(\alpha, \beta) = (1, 2)$ in the expression gives the Hernquist and NFW profiles, respectively.

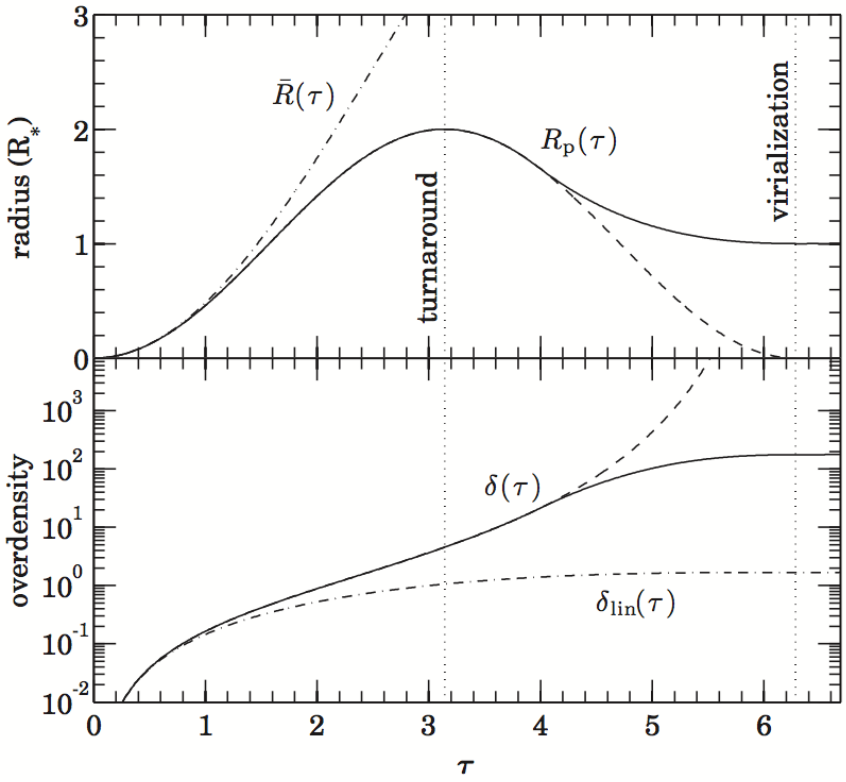


FIGURE 3.4 – Evolution of the spherical top-hat perturbation as a function of time. The growth of the radius r_t and the corresponding density δ_t of the spherical perturbation are shown by the solid line in the upper and lower panels respectively.

The NFW profile provides a better description of the density around virialized halos in numerical simulations [44]. It is completely characterized by the virial mass $M_{\rm vir}$ and the concentration parameter $c = R_{\rm vir}/R_s$ which is related to the characteristic over-density according to :

$$\delta_{\rm char} = \frac{\Delta_{\rm vir}\Omega_m}{3}\frac{c^3}{f(c)} \tag{3.54}$$

where $\Delta_{\rm vir} \simeq 178$ according to the spherical collapse model, [44] showed that the characteristic over-density is related to the halo's formation history, indeed halos that form earlier are more concentrated. The corresponding mass profile is thus given by :

$$M(R) = 4\pi\rho_{\rm crit}\delta_{\rm char}R_s^3 f(c) = M_{\rm vir}\frac{f(cR/R_s)}{f(c)} \tag{3.55}$$

A more recent updated version of the NFW profile has been proposed by [46] where the slope is a power-law function of the halo radius, such profile is called the Einasto :

$$\rho(R|M) = \rho_{-2}\exp\left[\frac{-2}{\alpha}\left(\left(\frac{R}{R_{-2}}\right)^{\alpha} - 1\right)\right] \tag{3.56}$$

where R_{-2} is the radius for which $\rho \propto R^{-2}$, for a NFW profile $R_{-2} = R_s$. The best-fit value for α typically spans the range $0.12 < \alpha < 0.25$ [47].

3.2.2.3 Number density of halos

At this stage, we know how to form halos from spherical collapse of a matter perturbation when the density is above a given threshold, $\delta_c = 1.69$ and how the matter is distributed within halos. We now need to know the halo mass function, i.e. the comoving number density of halos as a function of mass. A simple model for this was provided by Press and Schechter [48] who postulated that the probability $P(\delta_M > \delta_c)$ that $\delta_M > \delta_c(t)$ is the same as the mass fraction $F(> M, t)$ that is contained in halos with mass greater than M at time t, where δ_M is the filtered density field smoothed over a given mass M.

Assuming δ_M is a zero mean Gaussian random field with standard deviation σ_M, the probability that at a random point δ_M exceeds the threshold δ_c is :

$$P(\delta_M > \delta_c) = \frac{1}{\sqrt{2\pi}\sigma_M}\int_{\delta_c}^{\infty}\exp\left[-\frac{\delta_M^2}{2\sigma_M^2}\right]d\delta_M = \frac{1}{2}erfc\left[\frac{\delta_c}{2\sigma_M}\right] \tag{3.57}$$

where $erfc(x) = 1 - erf(x)$ with $erf(x)$ the error function. Thus, according to the above postulate, we can write

$$F(> M, t) = \frac{1}{2}erfc\left[\frac{\delta_c}{2\sigma_M}\right] \tag{3.58}$$

However, the PS postulate predicts that only half of all matter in the universe is enclosed in collapsed halos because only regions that are initially over-dense end up in collapsed structures. But under-dense regions can be enclosed within larger over-dense regions so in [48] they just added a factor two.

Now, if we define the mass function as $n(M, t)dM$ which is the number of halos with masses in the range $[M, M + dM]$ per comoving volume, it is related to the fraction of mass that is enclosed in halos with masses in the same range per comoving unit volume, so that :

$$n(M, t)dM = \frac{\bar{\rho}}{M}\frac{\partial F(> M, t)}{\partial M}dM \tag{3.59}$$

which yields

$$n(M,t)dM = 2\frac{\bar{\rho}}{M}\frac{\partial P(\delta_M > \delta_c)}{\partial M}dM \tag{3.60}$$

where we can use that :

$$\frac{\partial P(\delta_M > \delta_c)}{\partial M} = \frac{\partial P(\delta_M > \delta_c)}{\partial \sigma_M} \times |\frac{d\sigma_M}{dM}| \tag{3.61}$$

It thus yields :

$$n(M,t)dM = \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\frac{\delta_c}{\sigma_M}\exp\left[-\frac{\delta_M^2}{2\sigma_M^2}\right]\,|\frac{d\ln\sigma_M}{d\ln M}|\,dM \tag{3.62}$$

We often introduce the variable $\nu = \delta_c/\sigma_M$ which represents the peak height. The Press-Schechter mass function can thus be written :

$$n(M,t)dM = \frac{\bar{\rho}}{M^2}f_{PS}(\nu)|\frac{d\ln\nu}{d\ln M}|dM \;\text{ where }\; f_{PS}(\nu) = \sqrt{\frac{2}{\pi}}\nu e^{-\nu^2/2} \tag{3.63}$$

Although it provides very useful insights into the non-linear regime, the PS approach cannot be considered as a rigorous derivation, because of the multiplication by an additional factor of 2 and because it relies on the spherical top-hat model that explains the collapse of the dark matter halos as a spherical symmetric process while in reality halos can have more complicated shapes. Moreover, it does not take into account whether a halo of a given mass is included into a halo of larger mass. Despite these limitations, the comparison between the PS approach and numerical simulations shows that it gives roughly the right shape of the mass function and is correct up to an order of magnitude. Improvements to deal with the mentioned limitations have been proposed to statistically estimate how many small halos would reside in larger ones (the excursion set formalism, for more details see [49]). Further developments use simulation-calibrated formulas such as the empirical fit provided in [50] where they considered elliptical halos instead of spherical and which is usually referred as the Sheth-Tormen mass function (while the analytical paper came after by [51], or fitting formulas derived from simulations like in [52].

Figure 3.5 shows a comparison between numerical results (black) and the PS (dotted) and ST (dashed) mass functions where we can see that the PS prediction does not well reproduce the results for low-mass halos.

For what follows, we will use the PS and ST mass functions whose generic form is :

$$\nu f(\nu) \propto \left(1 + \frac{1}{(a\nu^2)^p}\right)\left(\frac{a\nu^2}{2}\right)^{1/2}\exp\left(-\frac{a\nu^2}{2}\right) \tag{3.64}$$

where $(a,p) = (1,2)$ and $(a,p) = (0.707, 0.3)$ correspond to the PS mass function and the ST mass function, respectively.

3.2.2.4 Halo clustering properties

The halo model approach, and more specifically the PS theory does not only open the door for an analytic calculation of the mean density of dark matter halos, it also provides insight into how these dark matter halos are correlated in space. At this stage, we have an expression of the abundance and spatial distribution of halos, as well as of the typical mass density profile around a halo. So we can now construct the two-point clustering properties of halos by assuming that the two-point correlation function has two contributions [19, 20],

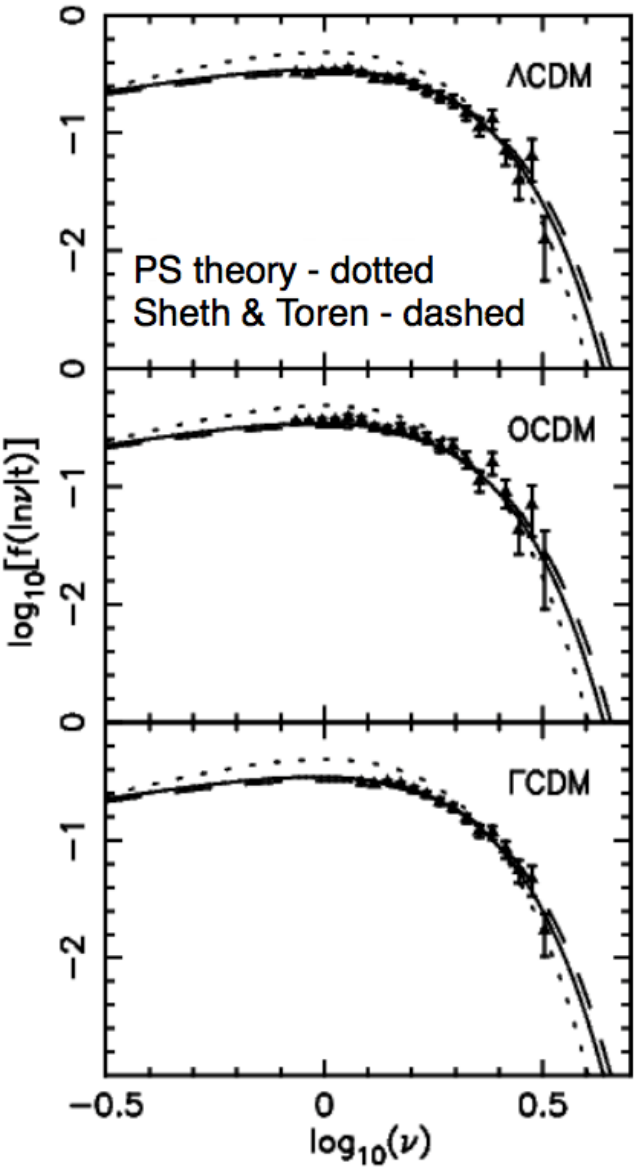


Figure 3.5 – PS and ST mass functions compared to results from numerical simulations for three cosmological models. The PS prediction does not well reproduce the results for low-mass halos.

a first one from the correlations between pairs within the same halo, denoted as ξ_{1h}, and a second one from the correlations between pairs of distinct halos, denoted as ξ_{2h} :

$$\xi_{hh}(r) = \xi_{1h}(r) + \xi_{2h}(r) \tag{3.65}$$

where $r = |\mathbf{x} - \mathbf{x}'|$, and each contribution can be expressed using the ingredients of the halo model that have been described previously. Assuming all mass is bound up into halos and that the halo density profile presented in the previous section depends on halo mass only, the density at a position $\mathbf{x}$ is obtained by summing up the contribution from each halo :

$$\rho(\mathbf{x}) = \sum_i m_i u(\mathbf{x} - \mathbf{x}_i | m_i) = \sum_i \int dm\, d^3x' \delta(m - m_i) \delta_{\mathbf{x}'-\mathbf{x}_i} m u(r|m) \tag{3.66}$$

where $u(r_i|m_i)$ is the radial number density profile of each halo (the density profile ρ divided by the total mass contained in the profile). The mean density $\bar{\rho}$ is thus defined by :

$$\bar{\rho} = \langle \rho(\mathbf{x}) \rangle = \int dm\, n(m)\, m \tag{3.67}$$

where $n(m)$ is the number density of halos of mass m. The auto-correlation of halos of mass M can thus be expressed in terms of ingredients of the halo model :

$$\xi_{hh}(r, M) = \langle \delta_h(\mathbf{x}, M) \delta_h(\mathbf{x}', M) \rangle = \frac{\langle n_h(\mathbf{x}, M) n_h(\mathbf{x}', M) \rangle}{\bar{n}_g^2(M)} - 1 \tag{3.68}$$

where each contribution, $\xi_{1h}(r)$ and $\xi_{2h}(r)$ can be expressed as functions of the ingredients of the halo model, the mass function which is related to $n(m)$ and the radial density profile within a halo $u(r|M)$.

Figure 3.6 shows the contribution from the 1-halo and 2-halo terms to the total power spectrum compared to the results obtained from N-body simulations. This technique has been refined by [53] by introducing fitting terms into the 2-halo and 1-halo contributions to the non-linear power spectrum obtained by comparison with simulations. This method, called 'Halofit', has been recalibrated by [54] using high-resolution N-body simulations.

3.3 The galaxy-halo connection

As one may have noticed, the theoretical framework presented in the previous sections has been developed for the clustering of dark matter, and also for dark matter halos using the halo model. Applications of its results to galaxy surveys is not trivial, because there is no guarantee that galaxies are faithful tracers of the dark matter field. On one hand, the development of numerical simulations allows us to resolve the dark matter structures and on the other hand, the emergence of large galaxy surveys allows us to identify large samples of galaxies. Thanks to both developments, it became possible to measure the spatial clustering properties and statistically connect these two distributions by inferring the connection between galaxies and halos.

3.3.1 Biased tracers

Halo formation is not a random process : halos are not a Poisson sampling of the matter field, instead they only form where the smoothed density field is above a given threshold given by the critical over-density for collapse. This is the *peak-bias picture* which

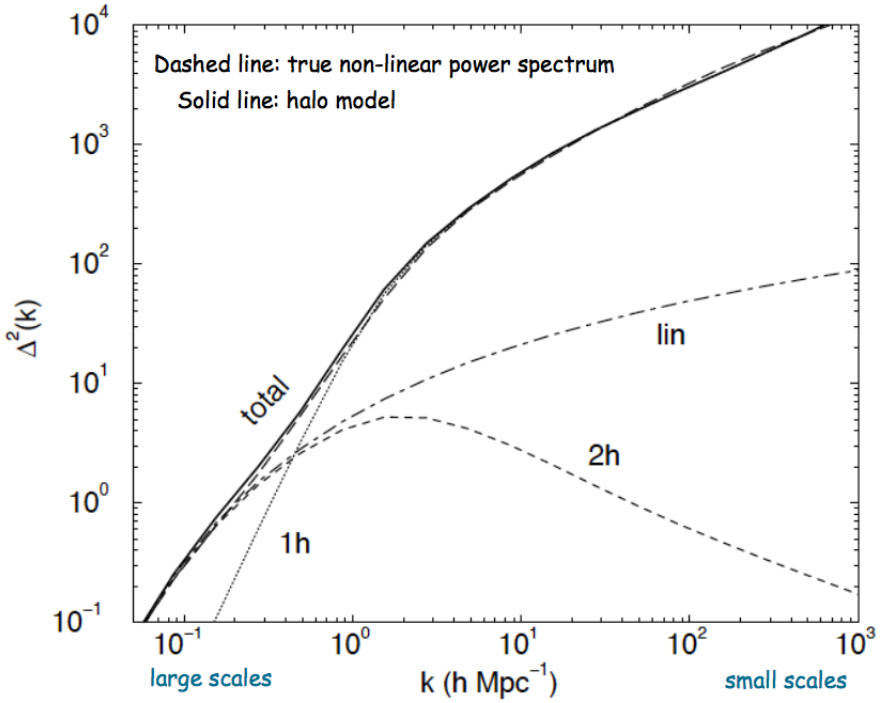


FIGURE 3.6 – 1-halo and 2-halo contributions to the total non-linear dimensionless power spectrum. From [21].

is illustrated by figure 3.7 where the threshold in the initial density field causes halos to be biased tracers of the mass distribution and thus

$$\frac{\delta N_g}{N_g} = b\frac{\delta\rho}{\rho} \tag{3.69}$$

This ground-breaking idea that actually galaxies do not trace faithfully the matter density field was first studied by Peebles, 1973 ; Tonry & Davis, 1979 and then Kaiser 1984 to explain the apparent correlation strength of Abell clusters with the idea that bright galaxies form only at the sites of high peaks in the initial density field. This approach has been extended to any object of a given mass by Bardeen et al. (1986,1996) and then by Cole & Kaiser (1989) using the *peak-background split model* (see also Mo & White 1996). In this framework, the perturbations in the initial dark matter density field is split into long wavelength contributions and short wavelength ones as illustrated in figure 3.8. The long wavelength parts should be much larger than the Lagrangian size of the halo, and it modulates the overall background fluctuations. The short wavelength parts are comparable or smaller than the halo size and they trigger local halo formation. In the presence of long wavelength perturbation, δ_l, which is rather constant at the scale of the halo, we can model its effect by modifying the halo formation density threshold, from $\delta_c(z)$ to $\delta_c(z) - \delta_l$. Therefore, the number density of galaxies $N_g(M)dM$ that reside within halos of mass M is therefore modulated :

$$N_g \longrightarrow N_g - \frac{dN_g}{d\delta_c}\delta_l \tag{3.70}$$

where Mo & White 1996 introduced the Lagrangian linear bias on large scales as $b^L = \frac{dN_g}{d\delta_c}$ and the Eulerian linear bias as $b^E = 1 + b^L$. The bias parameters can thus be obtained by differentiating the mass function, assuming a local biasing form.

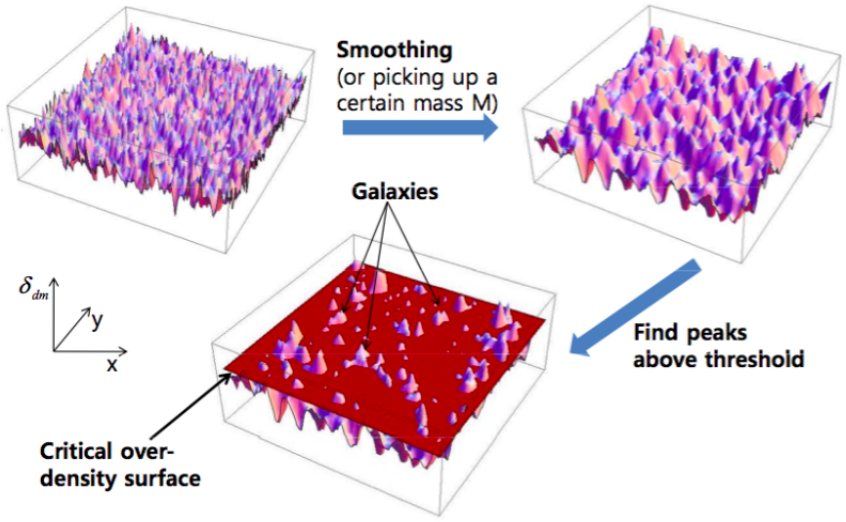


FIGURE 3.7 – Illustration of the peak-bias formalism : peaks above a given threshold in the initial density field correspond to regions where galaxies will form.

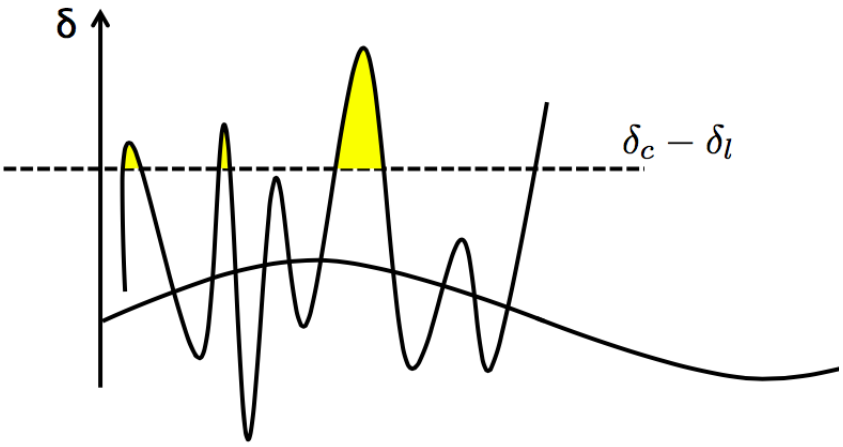


FIGURE 3.8 – In the peak-background split assumption, the initial perturbation is decomposed into a long wavelength and a short wavelength contributions.

General properties of the bias

When referring to the bias prescription, one should have in mind some general properties :

- **linear versus non-linear** : at very large scales ($> 80\,h^{-1}$Mpc) for Gaussian initial conditions, simulations show that one can reasonably assume that the relation between δ_g and δ_m is simply linear and the factor of proportionality defines the linear bias parameter. As we go to smaller scales, non-linearities start being non-negligible and the bias prescription becomes scale-dependent, especially for scales below $1\,h^{-1}$Mpc which corresponds to the transition between the 2-halo and the 1-halo term in the halo model.
- **local versus non-local** : by local, we mean that the bias relation depends only on the dark matter density field at the same position. At scales comparable to the size of halos, other degrees of freedom such as the tidal field in the vicinity become important and the biasing scheme is no longer considered as local. However, by smoothing the density field with a window radius larger than the size of halos, one can get rid of these small-scale effects and a locality is thus a good assumption.
- **deterministic versus stochastic** : if the bias were completely deterministic, the relation between the tracer and the matter density would present no additional scatter except the one expected from Poisson sampling of the underlying continuous density field which yields some noise. But, because of the finite size of halos, the halo exclusion effect contributes to the scatter at small scales. In addition, other contributions can arise from local environment properties which make the relation between halos and galaxies more complicated than just being mass-dependent only, with additional stochastic noise in the scatter between the galaxy and dark matter density fields.

The next section presents the approaches that have been developed to provide a theoretical modeling of the bias and the section after will describe the most important empirical models to connect the halo and galaxy properties.

3.3.2 Modeling the bias

Regardless of the way dark matter halos are populated to host the tracers of interest, more complicated theoretical frameworks have been developed to model more accurately the biasing scheme, taking into account non-linearities. The theoretical models are mostly based on the Lagrangian picture because it is easy to identify halos in the initial density fields but on the other hand, the Eulerian biasing models are more often used to interpret data because the ingredients of the Eulerian picture are closer to what we observe : density and velocity fields.

Eulerian bias model

One of the most popular bias model in Eulerian space consists in expanding the galaxy

density fluctuations $\delta_g(\mathbf{x})$ as a Taylor series [55] of the local matter density $\delta_m(\mathbf{x})$,

$$\delta_g(\mathbf{x}) = \sum_{n=1}^{\infty} \frac{b_n}{n!} \delta_m^n(\mathbf{x}) \tag{3.71}$$

The description given by equation 3.71 assumes that the galaxy bias is a function of the local matter density only. So this biasing scheme is local and deterministic. A non-local bias prescription was proposed by [56] which included a functional dependence of δ_g on the velocity divergence and gravitational potential, which are both non-local dependent on δ_m. It thus implies additional terms in the Taylor expansion to account for this non-locality dependence. The analysis of the DR14 quasar sample in Fourier space [57] uses the bias modeling of [56] with four bias parameters, b_1 the linear local bias parameter, b_2 a non-linear (second-order) bias parameter and b_{nl1}, b_{nl2} for non-local contributions.

Lagrangian bias model

This work relies on the local Lagrangian biasing scheme developed by [58] where locality of the bias in Lagrangian space is assumed and is not equivalent to locality in Eulerian space due to mass conservation. Indeed, let n^E (resp. n^L) denote the Eulerian (resp. Lagrangian) mass function, mass conservation implies : $n^E(M) = (1 + \delta^E)n^L$ where $1 + \delta^E$ accounts for the change in density in Eulerian space when the system collapses.

This assumption of locality in Lagrangian space implies that the Lagrangian density field $\rho^L_{\rm obj}(\mathbf{q})$ is assumed to be a function of a smoothed linear over-density at the same Lagrangian position :

$$\rho^L_{\rm obj}(\mathbf{q}) = \bar{\rho}_{\rm obj} F[\delta_R(\mathbf{q})] \tag{3.72}$$

where Matsubara (2008) introduced the Lagrangian bias function $F(\delta)$ and $\bar{\rho}_{\rm obj}$ is the comoving mean number density of the biased objects. The smoothed linear over-density can be expressed by :

$$\delta_R(\mathbf{q}) = \int d^3q' W_R(|\mathbf{q} - \mathbf{q}'|)\delta_l(\mathbf{q}') \tag{3.73}$$

where W_R is a smoothing kernel of size R and $\delta_l(\mathbf{q}')$ is the unsmoothed linear over-density.

Matsubara (2008) then defined the Lagrangian bias parameters $\langle F^{(n)} \rangle$ by :

$$\langle F^{(n)} \rangle = \frac{1}{\sqrt{(2\pi\sigma_R^2)}} \int d\delta \exp\left(-\frac{\delta^2}{2\sigma_R^2}\right) \frac{d^n F}{d\delta^n} \tag{3.74}$$

where σ_R^2 is the variance of the smoothed over-density field. Although more complicated, this expression shows that, here also, the bias parameters are obtained by differentiating the bias function $F(\delta)$. For a general biasing scheme, these parameters should be considered as additional (non-cosmological) parameters to be fitted by observations. Alternatively, one can also use a model for the bias function, such as the one derived from the halo model. Using the generic form for the PS and ST mass functions given by equation 3.64, Matsubara (2008b) derived the first two Lagrangian bias parameters :

$$\langle F' \rangle = \frac{1}{\delta_c(z)} \left[q\nu^2 - 1 + \frac{2p}{1 + (q\nu^2)^p} \right] \tag{3.75}$$

$$\langle F'' \rangle = \frac{1}{\delta_c^2(z)} \left[q^2\nu^4 - 3q\nu^2 + \frac{2p(2q\nu^2 + 2p - 1)}{1 + (q\nu^2)^p} \right] \tag{3.76}$$

In addition, by assuming the peak-background split model, we can find a relation between $\langle F' \rangle$ and $\langle F'' \rangle$ such that the only free parameter for the biasing model is the first Lagrangian one, $\langle F' \rangle$.

3.3.3 From halos to galaxies : empirical models

In section 3.2.2, we have shown how the halo model allows us to connect the matter clustering derived from theory to the statistical properties of halos. The ultimate connection that remains is the one between halo and galaxy properties. One can use the halo model described in the previous section to connect the statistical properties of halos with the ones of galaxies, by making the following change :

$$\frac{M}{\bar{\rho}} \longrightarrow \frac{\langle N \rangle_g}{\bar{N}_g} \tag{3.77}$$

$$\frac{M^2}{\bar{\rho}^2} \longrightarrow \frac{\langle N(N-1) \rangle_g}{\bar{N}_g^2} \tag{3.78}$$

$$u(r|M) \longrightarrow u_g(r|M) \tag{3.79}$$

where $\langle N_g \rangle$ describes the average number of galaxies that reside in a halo of mass M, $\bar{N}_g$ is the mean number density of those galaxies, and $u(r|M)$ (resp. $u_g(r|M)$) is the radial number density profile of halos (resp. galaxies in halos). In what follows, we will describe the most popular empirical models that are used to connect the halos properties to the galaxy ones. All these models are based on the assumption that one primary galaxy property is connected to one halo property and mass is a typical assumption for both.

3.3.3.1 Abundance matching

One of the simplest assumption one could make about the galaxy-halo connection is that the most massive galaxies reside in the most massive dark matter halos. This approach is generally called *abundance matching* and began by assuming that only one galaxy resides per halo [59, 60]. However, multiple distinct peaks in the density field within the radius of a dark matter halo correspond to smaller gravitationally-bound structures that orbit within the gravitational potential of the host halo, they are referred to as *subhalos*. Therefore, the abundance matching can be extended to sub-halos [61, 62], provided that the choice of the proxy to match halo (subhalo) properties with galaxy properties is appropriate. Such a proxy for galaxy can be the mass or luminosity for instance, and it is matched by abundance to the mass or velocity of the dark matter halo (subhalo) in which it lives. The challenge of such a prescription thus lies in the choice of the proxy, both for halos and galaxies. As highlighted in [62], this choice is even more complicated for subhalos whose identification is more difficult because of tidal stripping, i.e. when subhalos starts loosing large fraction of their mass up to even be tidally disrupted. For instance, [61] proposed to use the maximum circular velocity of halos $V_{\rm max}$ to match onto galaxies. Another possibility could be to use M_h at the time the subhalo was acrreted in order to provide a more robust primary property to stripping, or to introduce some scatter in the monotonic correspondence between halo and galaxy number densities (for more details, see [63].

Applications to quasars [34] modeled the clustering of the first year of eBOSS quasar data and produced high-fidelity quasar mock catalogues based on the BigMultiDark Planck simulation [35]. They used a modified (Sub)Halo Abundance Matching (SHAM) model to account for the specificities of the halo population hosting quasars based on the same approach developed in [64] for an ELG sample. This improvement consists in taking into account the incompleteness of the sample by splitting the probability function of selecting a halo into two terms corresponding to host and satellite halos. [64] used the virial mass of halos as primary property while [34] performed an HAM by using the

maximum circular velocity of halo as done in [65]. Their model relies on three parameters : the mean maximum circular velocity $V_{\rm mean}$, the width of the velocity distribution which is assumed to be Gaussian $\sigma_{\rm max}$ and the fraction of satellites $f_{\rm sat}$. These parameters are tuned to match the clustering of quasars that have been observed during the first year of eBOSS, however with the preliminary eBOSS data, they could not distinguish between models with different fractions of satellites. Moreover, as in [64], they found no dependency between $\sigma_{\rm max}$ and the clustering measurements. In fact, in the mass regime where quasars live, $\sigma_{\rm max}$ has an impact on the clustering at very small scales (below $0.5\,h^{-1}$Mpc) and thus it cannot be constrained with the current data. As a consequence, their model only depends on a single parameter, $V_{\rm mean}$ which is fixed by minimizing the χ^2 distribution when fitting the monopole of the correlation function between 10 and $40\,h^{-1}$Mpc. They found that quasars are hosted by halos with masses $\sim 10^{12.7} h^{-1} M_\odot$, which is an agreement with previous studies based on another empirical model described below, and their bias evolves from 1.54 ($z = 1.06$) to 3.15 ($z = 1.98$), which is also in agreement with previous measurements of the bias of quasars (a comparison of bias measurements including the one from this thesis work will be presented in section 5.4.3.

3.3.3.2 Halo occupancy distribution

Another popular way to describe the relationship between galaxies and dark matter halos is the Halo Occupation Distribution (HOD) framework which describes the probability distribution function $P(N \,|\, M)$ that a halo of mass M contains N galaxies of a given class. This prescription treats separately central and satellite galaxies and is fully characterized by its mean occupation number $\langle N|M\rangle$. Halo occupation statistics were first studied in relation with the halo model [20, 19, 66] in order to provide analytical expressions for the 1-halo and 2-halo terms as function of the probability for a central halo to host a galaxy $P(N_c \,|\, M)$, the probability for a satellite halo halo to host a galaxy $P(N_s \,|\, M)$ and the radial density profile of satellites within a halo.
Galaxy clustering can be used to constrain the HOD parameters. In fact, galaxy clustering at small scales $r < 1\,h^{-1}$Mpc is highly sensitive to the fraction of galaxies that are satellites because the number of pairs within a halo increases with the square of the number of satellites. Clustering at larger scales is most sensitive to the overall halo mass scale and to the scatter between halo and galaxy properties. The halo occupation framework predicts a transition scale in the clustering of galaxies at $r \sim 1$ Mpc that corresponds to the transition from the 1-halo term on smaller scales to the 2-halo term on larger scales.

Application to quasars Thanks to bigger and bigger quasar surveys, it has become possible to study quasar clustering in order to infer information on their astrophysical properties, in particular quasar populations show a high level of clustering from which we can infer that the occupy dark matter halos of mass $\sim 10^{12}\,h^{-1}.M_\odot$ at most epochs : [67] analysed the clustering of $\sim 4,000$ SDSS quasars, then [31] used a preliminary version of the SDSS-III BOSS sample with $\sim 27,000$ who showed that quasars have a duty cycle of about $\sim$1%. [67, 68] interpreted the quasar clustering within the HOD framework, but although their model matched the observed clustering, the parameters they found were unphysical because they did not take into account the duty cycle. [32] extended the analysis of the first sample of BOSS quasars [31] using more than $\sim 55,000$ BOSS quasars and showed that the statistics were not sufficient enough to detect a luminosity-dependence of the quasar properties.

In the analysis of the eBOSS DR14 quasar sample, a 5-parameter HOD model based on [69] was used in addition to a sixth parameter τ, which is used to model the duty cycle

of the quasars. In the approach we adopt, the number of satellites is independent of the presence of a central quasar and the expression of the HOD becomes :

$$\langle N_{\rm cen}\rangle_M = \tau \cdot \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm cen}}{\log \sigma_M}\right)\right] \tag{3.80}$$

$$\langle N_{\rm sat}\rangle_M = \left(\frac{M}{M_{\rm sat}}\right)^{\alpha_{\rm sat}} \cdot \exp\left(-\frac{M_{\rm cut}}{M}\right) \tag{3.81}$$

where $\langle N_{\rm cen}\rangle_M$ is the probability for a halo of mass M to host a central quasar and $\langle N_{\rm sat}\rangle_M$ is the number of satellite in a halo of mass M. The HOD parameters are tuned in order to match the peak of the mean density of quasars $n(z)$. Using the large-scale bias measurement of [33] where $b_Q = 2.45$, it also allows to determine the duty cycle of quasars, $\tau = 1.2\%$, which is in agreement with previous studies [31, 32]. The values of the parameters that reproduce the data are thus $M_{\rm cen} = 1.35\,10^{12}\,M_\odot$, $\log \sigma_M = 0.2$, $\alpha_{\rm sat} = 1$, $M_{\rm cut} = 10^8\,M_\odot$ and $M_{\rm sat} = 1.93 \times 10^{15}\,M_\odot$, which results in a population that consists of $\simeq 13\%$ satellites. Under these conditions the typical mass for dark matter halos hosting quasars is $M_{\rm cen} = 10^{12.5}\,M_\odot$.

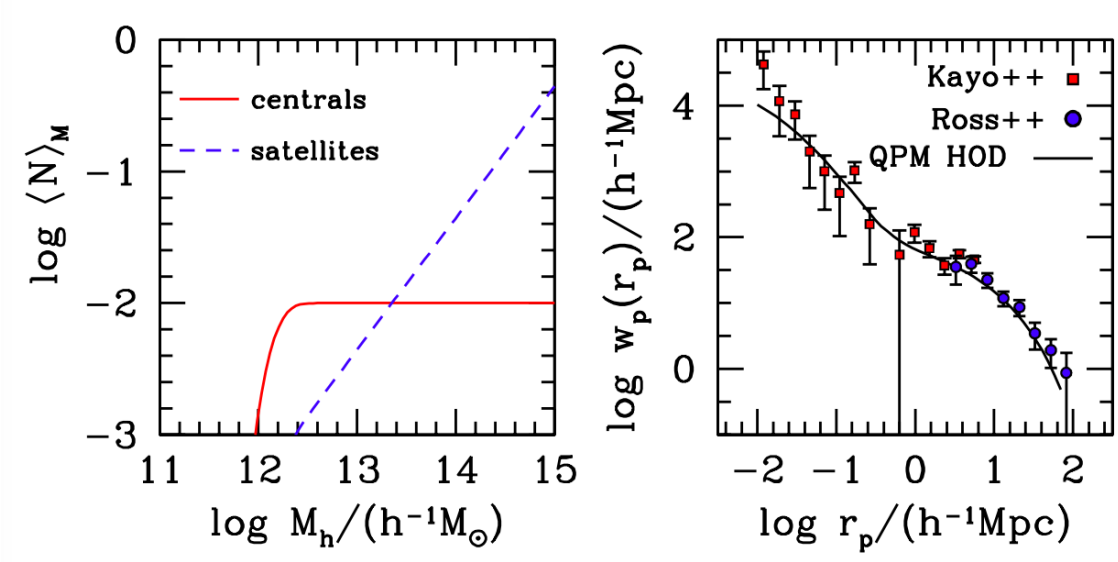


FIGURE 3.9 – Left panel : The mean number of central (red solid curve) and satellite (blue dashed curve) quasars per halo. The horizontal asymptote of 0.01 central quasars per halo reflects the adopted duty cycle of 1% for quasars in the eBOSS sample. Left panel : The projected correlation function obtained from the HOD parameters used in this analysis (black solid curve) compared to measurements of quasars from [70] in blue circles and of [71] in red squares.

The left panel of figure 3.9 shows the HOD model with the above parameters that match both the peak of $n(z)$ and the large-scale bias measured by [33] which fixes the satellite fraction. The right panel compares the projected correlation function predicted by the HOD model to measurements of quasar clustering from [70] using a sample of SDSS quasars and from [71] using a sample of 26 binary quasars at $0.6 < z < 2.2$ from the Sloan Quasar Lens Search (SQLS).

Although SHAM and HOD models can provide a good agreement with the observed quasar clustering, it is still not clear to what extend such standard prescriptions can

describe rare and heavily biased quasars. Moreover, the halo properties are still studied to characterize the evolution the duty cycle with luminosity and redshift [for a recent study using eBOSS quasars see 33]. Other studies showed that the formation of such quasars could also be the result of a major merging of gas-rich galaxies [72, 73, 74] . Precise measurements of quasar clustering using the final eBOSS sample would start being powerful enough to detect whether quasar clustering is luminosity-dependent and would thus discriminate between different scenarios.

3.3.3.3 Secondary properties

The two frameworks that have been described previously relies on the assumption that the relation between galaxies and halos is set by one halo property to be matched to one galaxy property, but there is no a priori reason to limit the prescription to a single primary property. As already mentioned, other dependencies can lead to a non-monotonic relation between the galaxy and dark matter halo density fields and halos and galaxies at fixed mass can have a different history. In fact, halos that assemble earlier are found to be more strongly clustered than halos of the same mass that assemble later. The effect according to which the clustering of halos at fixed mass depends on other properties than M_h is called *halo assembly bias*. The effect also exists for galaxy clustering and is thus referred to *galaxy assembly bias*. We refer the reader to the recent comprehensive review of the galaxy-halo connection [75] for more details on the prescriptions beyond the mass-only ansatz. In what follows, I briefly list the possible secondary properties that could be used to characterize the halo assembly bias :

— peculiar velocity, also referred to as velocity bias : the assumptions according to which the radial distribution of satellite galaxies inside DM haloes follows that of mass and that central galaxies are defined as at rest at the centre of their DM halo make galaxy velocities unbiased with respect to the mass velocity field. However, there are some observational evidence that galaxies do not exactly follow the same radial distribution as DM and exhibit some velocity bias. Neglecting them yields to an underestimate of f by 1-3 per cent depending on redshift [76].

— halo concentration : each halo at fixed mass can be divided into sub-categories depending on its concentration, this additional property has been implemented as an extension of the HOD model with two halo properties, dubbed *decorated HOD* [77].

— distance to halo center : the spatial distribution of satellites within halos will be different depending on how close to the center satellite halos are.

It is worth noting that all these empirical models rely on the ability of numerical simulations and halo finders to track and resolve halos and their substructures in order to make proper comparisons between the proxy one choose for galaxy and the one for halos. In addition, even the definition of halo radius is under debate, as the common one using the virial radius R_{vir} may not be the most physically motivated definition of the boundary of a dark matter halo. [78] have emphasized that it could lead to unphysical interpretations about halo mass accretion histories, as measuring halo growth using M_{vir} for instance would imply to infer higher halo growth than expected just because of the halo boundary being defined to larger radii with time. Several authors have suggested another definition, such as the halo mass M_h at the time the subhalo was acrreted which is more robust to stripping [61] or the *splashback radius* which specifies the radius at which matter that is bound to the halo can orbit after first collapse [79, 80]. It provides a more physically motivated halo boundary separating the infall and multi-stream regions.

3.4 Real-to-redshift space mapping

A third origin for non-linearities arises from our method to infer distances from spectroscopic surveys. As explained in the introduction, large-scale structure maps produced by estimating distances from redshifts reveal an anisotropic distribution in redshift space. If the Universe were statistically isotropic and homogeneous on large scale, these redshifts would accurately measure radial distance from the observer, so that the mapping between real space (r-space) to redshift space (s-space) would be an identity. It instead exhibits an anisotropy with respect to the line-of-sight (LOS) direction. These differences arise because galaxy recession velocities include components from both the Hubble flow and peculiar velocities associated with any inhomogeneous structure.

The redshift-space comoving position $\mathbf{s}$ of a galaxy differs from its real-space comoving position $\mathbf{x}$ due to its comoving peculiar velocity $\mathbf{u}$,

$$\mathbf{s} = \mathbf{r} + u_z(\mathbf{r})\hat{\mathbf{z}} \tag{3.82}$$

where I have taken the LOS to be the z-axis, $\hat{\mathbf{z}}$ is the unitary vector, u_z is the LOS component of the peculiar velocity $\mathbf{u}$ [2]. We adopt the plane-parallel approximation, so the LOS direction is fixed for all objects.
In what follows, we consider objects which are not part of a larger collapsed system, like individual galaxies in clusters. Thus the peculiar velocitiers are usually small, of the order of $10^{-3}\,\mathrm{km.s^{-1}}$ or less.
In this section, we present the most commonly used approaches in Fourier space and configuration space to model this real-to-redshift space mapping due to the LOS component of peculiar velocities when measuring redshifts.

3.4.1 RSD in the linear regime

The redshift space $\mathbf{s}$ is a distorted version of the real space $\mathbf{r}$. We require the mapping to be one-to-one so that these volume elements contain the same galaxies. The exact Jacobian for the real-space to redshift-space transformation is thus obtained by imposing mass conservation between the over-density in redshift space δ_t^s and the one in real space δ_t where the subscript t accounts for the density field of a biased tracer, either galaxy or quasar,

$$(1+\delta_t^s)d^3\mathbf{s} = (1+\delta_t)d^3\mathbf{r} \tag{3.83}$$

The redshift space distortions only affect the LOS direction and in the distant-observer limit we neglect the curvature of the sky so,

$$\frac{d^3\mathbf{s}}{d^3\mathbf{r}} = \frac{s^2}{z^2}\frac{ds}{dz} \tag{3.84}$$

with $s = z + u_z$ in the LOS direction (z-axis) The Jacobian thus becomes :

$$\frac{d^3\mathbf{s}}{d^3\mathbf{r}} = \left(1+\frac{u_z}{z}\right)^2\left(1+\frac{\partial u_z}{\partial z}\right) \tag{3.85}$$

so that

$$1+\delta_t^s = (1+\delta_t)\left(1+\frac{u_z}{z}\right)^{-2}\left(1+\frac{\partial u_z}{\partial z}\right)^{-1} \tag{3.86}$$

2. In this notation, the peculiar velocity $\mathbf{u}$ is taken in units of h^{-1}Mpc and the conversion to $\mathrm{km.s^{-1}}$ is given by the Hubble parameter : $u[h^{-1}\mathrm{Mpc}] = u[\mathrm{km.s^{-1}}]/aH$

We make the **linear perturbation theory approximation**, i.e. perturbations and their gradients are small : $\delta, \mathbf{u}, \nabla\delta, \nabla\mathbf{u} \ll 1$ so that products of these quantities are second-order, hence small and we cand drop them. We also make a separate approximation, the **distant-observer limit**, assuming that $u_z \ll z$, i.e. we consider sufficiently large distances that the peculiar velocities are small compared to the Hubble recession velocity.
According to these assumptions, equation 3.86 simplifies into

$$1 + \delta_t^s = (1 + \delta_t)\left(1 + \frac{\partial u_z}{\partial z}\right)^{-1} \tag{3.87}$$

which gives at linear order :

$$\delta_t^s = \delta_t - \frac{\partial u_z}{\partial z} \tag{3.88}$$

In Fourier space, the second term on the right gives $-ik_z u(k)$ where $u(k)$ is the Fourier transform of the tracer velocity field u. In the linear regime we can assume the velocity field is irrotational as long as there is one single flow only. It is therefore completely described by its divergence $\theta = \nabla \cdot \mathbf{u}$. In Fourier space, it gives :

$$u_z(k) = i\frac{k_z}{\mathbf{k}}\theta(k) \tag{3.89}$$

It thus gives for the tracer over-density in redshift space :

$$\delta_t^s(k) = \delta_t(k) + \left(\frac{k_z}{\mathbf{k}}\right)^2 \theta(k) = \delta_t(k) + \mu_k^2\theta(k) \tag{3.90}$$

where $(k_z/\mathbf{k})^2$ to the cosine of the angle between the LOS and the pair separation vector, denoted μ_k here in Fourier space. Then its power spectrum defined by $P_t^s(k) = \langle|\delta_t^s(k)|^2\rangle$:

$$P_t^s(k) = \langle|\delta_t(k)|^2 + 2\mu_k^2\langle\delta_t(k)\theta_t(k)\rangle + \mu_k^4\langle|\theta_t(k)|^2\rangle \tag{3.91}$$

$$= P_{\delta\delta,t} + 2\mu_k^2 P_{\delta\theta,t} + \mu_k^4 P_{\theta\theta,t} \tag{3.92}$$

where we define the density, density-velocity and velocity-velocity power spectra, $P_{\delta\delta,t}$, $P_{\delta\theta,t}$ and $P_{\theta\theta,t}$.

As explained in section 3.3.1, the galaxy distribution differs from the mass distribution by what we call the linear bias. On large scales, in the linear regime, one expects the bias to be a constant multiplicative factor to the mass density field $\delta_t = b\delta_m$. The linear regime also implies that there is a **linear coupling** between the velocity and density fields so that $\theta_m = f\delta_m$. In addition, we assume that the radial distribution of satellite galaxies inside the DM halos follows that of mass and that central galaxies are defined to be at rest at the centre of their DM halo. According to these assumptions, galaxies velocities are considered to be unbiased w.r.t the mass velocity field so we neglect any velocity bias and $\theta_m = \theta_t$. In the linear regime, $P_{\delta\delta,m} = P_{\delta\theta,m} = P_{\theta\theta,m} = P_m$. It thus gives the *Kaiser formula* [81] for the power spectrum in redshift space :

$$P_t^s(k) = (b^2 + 2fb\mu_k^2 + f^2\mu^4)P_m(k) \tag{3.93}$$

The Kaiser effect thus refers to the apparent enhancement of the clustering along the LOS at large scales.

In this thesis, we work in configuration space so we give the expression for the first even Legendre multipoles in redshift space in linear theory :

$$\xi_0^s = (b^2 + \frac{2}{3}fb + \frac{1}{5}f^2)\xi_m^r(r) \tag{3.94}$$

$$\xi_2^s = (\frac{4}{3}bf + \frac{4}{7}f^2)[\xi_m^r(r) - \xi_m'(r)] \quad \text{where} \quad \xi_m'(r) = \frac{3}{r^{-3}}\int_0^r \xi_m(r')r'^2 dr' \tag{3.95}$$

$$\xi_4^s = \frac{8}{35}f^2\xi_m''(r) \quad \text{where} \quad \xi_m''(r) = \frac{5}{r^5}\int_0^r \xi_m(r')r'^4 dr' \tag{3.96}$$

Dispersion model So far, we have neglected the motion of galaxies inside virialized dark matter halos that give rise to the Fingers-of-God effect which suppress power at high k (it is sometimes called the damping effect) or small scales s. A phenomenological approach consists in treating the Kaiser enhancement at large scales and the FOG damping at small scales as two independent effects so that both can be factorized as follows :

$$P^s_{\text{pheno}}(k) = D_{\text{FOG}}(k\mu f\sigma_\nu)P^s_{\text{Kaiser}}(k,\mu) \tag{3.97}$$

where $P^s_{\text{Kaiser}}(k,\mu)$ is given by equation 3.93. This expressions corresponds to the dispersion model, for limitations of the factorization see [82]. [83] first modeled the FOG effect as a Gaussian noise so that :

$$P^s_t(k) = e^{-k^2\mu^2\sigma_v^2}(b + f\mu^2)P_m(k) \tag{3.98}$$

where σ_v corresponds to the dispersion velocity along the line-of-sight. This parameter is usually treated as a free parameter as the physics involved at these small-scales is highly non-linear which makes the modeling of the dispersion velocity distribution very challenging. In particular, the amplitude of the effect is strongly dependent on the mean halo mass and satellite fraction of the population under consideration. A Lorentzian prescription for the FoG effect is also often used so in general we can have :

$$D_{FoG}(k,\mu,\sigma_v) = \begin{cases} e^{-k^2\mu^2\sigma_v^2} & \text{for Gaussian} \\ \frac{1}{1+k^2\mu^2\sigma_v^2/2} & \text{for Lorentzian} \end{cases} \tag{3.99}$$

3.4.2 RSD in the non-linear regime

In the non-linear regime, [82] extended the Kaiser formula so that equation 3.93 becomes :

$$P^s_t(k) = b^2P_{\delta\delta,m}(k) + 2fb\mu_k^2P_{\delta\theta,m}(k) + f^2\mu^4P_{\theta\theta,m}(k) \tag{3.100}$$

where $P_{\delta\delta,m}$, $P_{\delta\theta,m}$ and $P_{\theta\theta,m}$ are computed using Eulerian perturbation theory techniques, i.e RPT [12, 13, 18], RegPT [15] or gRPT (Crocce, Blas and Scoccimarro in prep).

Equation 3.100 is then injected in the dispersion model given by equation 3.98 to include the FOG effect. However, comparisons to N-body simulations showed that the model given in equation 3.98 can differ by several percents, even when including the non-linear expression of $P_m(k)$. In particular, [84] showed that two additional terms appeared in the redshift-space power spectrum when using the dispersion model that are due to the coupling between density and velocity fields associated with Kaiser and FOG effects. The corrected dispersion model, denoted as the TNS model in what fallows, can thus be written as :

$$P_t(k,\mu) = e^{-k^2\mu^2\sigma_v^2}\left[b^2P_{\delta\delta,m}(k) + 2fb\mu_k^2P_{\delta\theta,m}(k) + f^2\mu^4P_{\theta\theta,m}(k) + b^3A(k,\mu,f) + b^4B(k,\mu,f)\right] \tag{3.101}$$

where :

$$A(k,\mu,f) = (k\mu f)\int \frac{d^3q}{(2\pi)^3}\frac{q_z}{q^2}[B_\sigma(\mathbf{q},\mathbf{k}-\mathbf{q},-\mathbf{k}) - B_\sigma(\mathbf{q},\mathbf{k},-\mathbf{k}-\mathbf{q})] \tag{3.102}$$

$$B(k,\mu,f) = (k\mu f)^2\int \frac{d^3q}{(2\pi)^3}F(\mathbf{q})F(\mathbf{k}-\mathbf{q}) \tag{3.103}$$

where $F(\mathbf{q})$ is a combination of $P_{\delta\theta}$ and $P_{\theta\theta}$ while B_σ corresponds to the cross bispectra that encode the 3-point correlations between the velocity and density fields, i.e. $\langle\theta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle$, $\langle\theta(\mathbf{k}_1)\delta(\mathbf{k}_2)\theta(\mathbf{k}_3)\rangle$, $\langle\theta(\mathbf{k}_1)\theta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle$, ...

In its original paper, the TNS model was proposed using the 1-loop SPT with a linear bias. Extensions to the TNS model include other PT codes to compute $P_{\delta\delta,m}$, $P_{\delta\theta,m}$ and $P_{\theta\theta,m}$ (2-loop SPT, RegPT, RPT, gRPT) and non-local Eulerian bias (e.g. [56, 85]). Comparisons in Fourier and redshift space between the TNS model and the original dispersion model show that the TNS performances are much better and can reach the percent accuracy at $k_{\rm max} = 0.10 - 0.15\,h{\rm Mpc}^{-1}$ at $z = 0$ and $k_{\rm max} = 0.20\,h{\rm Mpc}^{-1}$ at $z = 1$ [18]. Although the TNS model has been developed in Fourier space, it can be easily applied to configuration space observables.

Streaming model Another approach to model the redshift-space correlation function has been developed by [3] first to estimate the galaxy pair-wise velocity dispersion from the suppression of the correlation on small scales. It was thus extended to large scale [86, 87], then improved [82] and updated to its actual form [88]. In this approach, the redshift space correlation function can be seen as a convolution of the real-space correlation function with the pair-wise velocity probability distribution function $\mathcal{G}$:

$$1 + \xi^s(s_\perp, s_\parallel) = \int d^3r_\parallel [1 + \xi^r(r)]\mathcal{G}(s_\parallel - r_\parallel, v_{12}, \sigma_{12}) \tag{3.104}$$

where $\mathcal{G}$ maps the pairs at separation $r_\parallel$ to separation $s_\parallel$ due to peculiar velocities with probability F. The challenge is thus to model this PDF in terms of the linear power spectrum, cosmological parameters and tracer bias. Near its peak, $\mathcal{G}$ can be approximated by a Gaussian centered on $\mu v_{12}(r)$ [82] and in order to recover a streaming model that agrees on large scales with N-body simulations, one needs to assume a Gaussian distribution with non zero pair-wise velocity distribution and a scale-dependent pair-wise velocity dispersion [88] so that $\mathcal{G}$ becomes :

$$\mathcal{G} = \frac{1}{\sqrt{2\pi\sigma_{12}^2(r,\mu)}} \exp[\frac{(s_\parallel - r_\parallel - \mu v_{12})^2}{2\sigma_{12}^2(r,\mu)}] \tag{3.105}$$

The non-zero pair-wise velocities come from the coupling between the density and velocity fields on large scales which gives rise to a coherent infall velocity v_{12} between pairs of matter tracers along the line of separation $\hat{\mathbf{r}}$. Note that this coherent infall is an average of the tracer velocity flow, which can be seen as a number-weighted average in the sense that galaxies are preferentially located in high-density regions, associated with a probability $\propto (1 + \delta)$. So we need to include this probability of finding a galaxy at a given position when defining v_{12} :

$$\begin{aligned} v_{12}(r) &= \hat{\mathbf{r}}\langle\mathbf{v}_{12}(\mathbf{r})\rangle & (3.106)\\ &= \hat{\mathbf{r}}\langle[\mathbf{v}(\mathbf{x}+\mathbf{r}) - \mathbf{v}(\mathbf{x}][1+\delta(\mathbf{x})][1+\delta(\mathbf{x}+\mathbf{r})]\rangle & (3.107)\\ &= \hat{\mathbf{r}}[\langle[\mathbf{v}(\mathbf{x}+\mathbf{r})\delta(\mathbf{x})\rangle - \langle\mathbf{v}(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle] + \text{higher order terms} & (3.108)\end{aligned}$$

where if we assume δ and $\mathbf{v}$ are Gaussian fields, the density and velocity at the same point are uncorrelated.

Using the linear coupling $\theta = f\delta$ and taking the Fourier transform of δ and $\mathbf{v}$ gives :

$$v_{12}(r) = ifb \int \frac{d^3\mathbf{k}}{(2\pi)^3}\mathbf{k}\cdot\mathbf{r}e^{-i\mathbf{k}\cdot\mathbf{r}}P_m(k) \tag{3.109}$$

In the isotropic case, $d^3\mathbf{k} = k^2 dk \sin\theta d\theta d\phi$ and $\mathbf{k}\cdot\mathbf{r} = kr\cos\theta$. When integrating w.r.t θ, one can make a change of variable $X = \cos\theta$ so that the following integral appears which gives the first-order spherical Bessel function j_1 :

$$\int_0^1 dX e^{ikrX} 2X = -2ij_1(kr) \tag{3.110}$$

$$v_{12}(r) = -r\frac{fb}{2\pi^2}\int kP_m^r(k)j_1(kr)dk \tag{3.111}$$

One can also obtain an expression for σ_{12} from linear theory predictions [87, 88] :

$$\sigma_{12}^2(r,\mu^2) = 2[\sigma_v^2 - \mu^2\Psi_\parallel(r) - (1-\mu^2)\Psi_\perp] \tag{3.112}$$

where $\sigma_v^2 = \frac{1}{3}\langle v(\mathbf{x})v(\mathbf{x})\rangle$ is the one-dimensional velocity dispersion and

$$\Psi_\parallel(r) = \frac{f^2}{2\pi^2}\int dkP_m(k)\left[j_0(kr) - \frac{2j_1(kr)}{kr}\right] \tag{3.113}$$

$$\Psi_\perp(r) = \frac{f^2}{2\pi^2}\int dkP_m(k)\frac{j_1(kr)}{kr} \tag{3.114}$$

with j_0 and j_1 the spherical Bessel functions.
The streaming approach is therefore based on three main ingredients : the real space correlation function $\xi(r)$, the mean infall pair-wise velocity $v_{12}(r)$ and the pair-wise dispersion velocity σ_{12}. As explained in section 3.1, linear theory breaks down in the quasi-linear regime due to the non-linear evolution of the density and velocity fields that starts to be non-negligible and as mentioned above in the section, [82] explored the corrections induced in the power spectrum due to the non-linear evolution of the density and velocity fields and calculated using perturbation linear theory. In the streaming model, it means that more complicated expressions from perturbation theory models will replace the ones from linear theory for $\xi(r)$, v_{12} and σ_{12}.

In its original form, the GS model was used in association with SPT to compute $\xi(r)$, v_{12} and σ_{12} at intermediate scales $(20 - 60\,h^{-1}\text{Mpc})$ [88, 89]. In this thesis, we use the Convolution Lagrangian Perturbation Theory (CLPT) developed in [90] and extended in [91], which improves the work done by [17] who introduced a perturbative scheme which is different from the standard one given by 3.46. The perturbative approach allows us to reach smaller scales than the ones described by linear theory. Moreover, working in Lagrangian space makes the inclusion of RSD and the bias easier. The bias model is based one local Lagrangian biasing scheme developed in [58] and presented in section 3.3.2. The density field of a biased tracer $\delta_t(\mathbf{x},\mathbf{t})$ is assumed to be a function F of a smoothed linear over-density at the same Lagrangian position : $\delta_R(\mathbf{q})$

$$1 + \delta_t(\mathbf{x},\mathbf{t}) = \int d^3qF[\delta_R(\mathbf{q})]\delta_D[\mathbf{x} - \mathbf{q} - \mathbf{\Psi}(\mathbf{q},t)] \tag{3.115}$$

where the expression for $\delta_R(\mathbf{q})$ is given by equation 3.73 and $\mathbf{x}(\mathbf{q},t) = \mathbf{q} + \mathbf{\Psi}(\mathbf{q},t)$ (equation 3.31) is the transformation from Eulerian to Lagrangian coordinates which defines the displacement field $\mathbf{\Psi}$.

Then, [90] derived the exact configuration space expression which is anologue of the derivation by [17] in Fourier space but the difference lies in the choice of expansion which can be seen as a partial resummation of the result of [17, 58]. By statistical homogeneity,

the two-point correlation function $\xi_t(\mathbf{r}) = \langle \delta_t(\mathbf{x}_1)\delta_t(\mathbf{x}_2)\rangle$ depends only on the difference in Lagrangian coordinates $\mathbf{q} = \mathbf{q}_2 - \mathbf{q}_1$. It can be written as :

$$1 + \xi_t(r) = \int d^3q \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{q}-\mathbf{r})} \int \frac{d\lambda_1}{2\pi}\frac{d\lambda_2}{2\pi} \tilde{F}_1\tilde{F}_2 K(\mathbf{q},\mathbf{k},\lambda_1,\lambda_2) \tag{3.116}$$

where K is defined by :

$$K(\mathbf{q},\mathbf{k},\lambda_1,\lambda_2) = \langle e^{i(\lambda_1\delta_1+\lambda_2\delta_2+\mathbf{k}\cdot\Delta)}\rangle \tag{3.117}$$

with $\Delta = \mathbf{\Psi}_2 - \mathbf{\Psi}_1$ and λ_1 (resp. λ_2) is the Fourier transform of the bias parameter obtained from $F(\delta)$ for object 1 (resp. object 2).
[17] showed that applying the cumulant expansion theorem allows us to expand the expectation value in terms of cumulants :

$$\langle e^{iX}\rangle = \exp\left[\sum_{N=1}^{\infty} \frac{i^N}{N\,!}\langle X^N\rangle_c\right] \tag{3.118}$$

where $\langle X^N\rangle_c$ denotes the N^{th} cumulant of the random variable X. Given that we deal with Gaussian fields (Ψ is a function of δ_l which is assumed to be Gaussian), only the second cumulant survives ($N = 2$) which gives :

$$\langle e^{i(\lambda_1\delta_1+\lambda_2\delta_2+\mathbf{k}\cdot\Delta)}\rangle = \exp[\langle(\lambda_1\delta_1 + \lambda_2\delta_2 + \mathbf{k}\cdot\Delta)^2\rangle_c] \tag{3.119}$$

$$= \exp[(\lambda_1^2+\lambda_2^2)\sigma_R^2 + A_{ij}k_ik_j + 2\lambda_1\lambda_2\xi_R + 2(\lambda_1+\lambda_2)U_ik_i] \tag{3.120}$$

where σ_R^2 is the variance of the smoothed linear density field $\delta_{1,2}$, $\xi_R = \langle\delta_1\delta_2\rangle_c$ is the corresponding smoothed linear correlation function, A_{ij} is related to the velocity-velocity coupling and U_i is related to the density-velocity coupling. Then, [90] performed a partial expansion noting that some terms (σ_R^2 and A_{ij}) in the above expression did not vanish in the large-scale limiti when $|\mathbf{q}| \to \infty$ and should thus be kept exponentiated instead of taking them into account in the expansion. To get $\xi(r)$, we integrate equation 3.120 over λ_1, λ_2, $\mathbf{k}$ and $\mathbf{q}$. The integrations w.r.t. λ_1, λ_2 introduce the Lagrangian bias parameters through the idendity [58] :

$$\int \frac{d\lambda}{2\pi}\tilde{F}(\lambda)(i\lambda)^n e^{-\lambda^2\sigma_R^2/2} = \langle F^{(n)}\rangle \tag{3.121}$$

where we consider the first two Lagrangian parameters, F' and F''.

Despite the fact that CLPT dramatically improves the description of correlation function in real space, it remains inaccurate on quasi-linear scales for the quadrupole in redshift space. To overcome this deficiency, [91] extended the formalism to include the calculation of velocity moment statistics such as the pairwise infall velocity v_{12} and pairwise velocity dispersion σ_{12}. These two ingredients with the correlation function in real space $\xi(r)$ are used as inputs in a Gaussian-Streaming (GS) model proposed by [88] and reviewed in section **??**. The relative peculiar velocity between two tracers at Eulerian coordinates $\mathbf{x}_1$ and $\mathbf{x}_2$ is related to the time derivative of the difference between the displacement fields, so that $\mathbf{v}(\mathbf{x}_2) - \mathbf{v}(\mathbf{x}_1) = \dot{\Delta}/H$ (where $\mathbf{v}$ is in unit of h^{-1}.Mpc). The term K defined in equation 3.117 thus contains an additional term which is proportional to $\dot{\Delta}/H$. After integrating w.r.t λ_1, λ_2, $\mathbf{k}$ and $\mathbf{q}$, [91] showed that the CLPT prediction for velocity statistics can be written in the same manner as the prediction for the clustering found in [90], i.e. as function of the first two Lagrangian parameters $\langle F'\rangle$ and $\langle F''\rangle$ (whose expressions are given by equations 3.76) :

$$A = A_0 + \langle F'\rangle A_{10} + \langle F''\rangle A_{01} + \langle F'\rangle^2 A_{20} + \langle F''\rangle^2 A_{02} + \langle F'\rangle\langle F''\rangle A_{11} \tag{3.122}$$

TABLE 3.2 – The different RSD models that have been applied to BOSS and eBOSS data.

observable	name	theory	BOSS analysis	eBOSS analysis
$P(k)$	eTNS+RPT	[12, 13, 18]	[94, 95, 96]	[57]
$P(k)$	eTNS+RegPT	[15]	[97, 98]	–
$P(k)$ / $\xi(s)$	eTNS+gRPT	Blas, Crocce, Scoccimarro in prep	[99, 100]	[101]
$\xi(s)$	GS+SPT	[88]	[89, 102, 103]	–
$\xi(s)$	GS+CLPT	[90, 91]	[104, 105]	[106]

where A can be replaced by $\xi(r)$, v_{12}, $\sigma_{12,\parallel}$ or $\sigma_{12,\perp}$ and each A_{ij} coefficient is calculated in CLPT. The terms A_{10} and A_{01} correspond to corrections at 1^{st}-order ; A_{20}, A_{02} and A_{11} at 2^{nd}-order. In what follows, for simplicity we will refer to the Lagrangian parameters as F' and F''. [91] showed that the second bias parameter F'' has a negligle effect on the correlation function, only affecting small scales (below $\sim 20\,h^{-1}$Mpc). So here, it is fixed under the peak-background split assumption [92] using the Sheth-Tormen mass function [ST, 50]. We will show in section 4.2.3 the effect on the cosmological parameters when using the Press-Schetcher mass function (PS, [48]) or when setting F'' as a free parameter when fitting on observables in redshift space.

Comparisons in configuration and redshift space of different RSD models for the analysis of the BOSS final sample can be found in [93]. Table 3.2 summarizes the different RSD models that have been applied to BOSS and eBOSS data already.

3.4.3 Beyond the Gaussian pair-wise velocity PDF

[82] also showed that the PDF is strongly non-Gaussian at all scales and that actually there is no one single PDF involved in equation 6.10 but rather an infinite number corresponding to different scales and angles of the velocities with respect to the line joining the pair. More recently, this non-Gaussian PDF has also been highlighted by [76] who showed that there is not any simple general functional form for the pairwise velocity PDF that matches all scales and for all types of tracers.

The full pair-wise velocity PDF of halo pairs is calculated at different pair separations using the BigMDPL simulation [3] showed in the top panel of figure 3.10. The presence of skewness and exponential tails at all scales clearly show that the Gaussian assumption made in the streaming model according to which $\mathcal{G}$ only depends on its first two moments is not satisfactory. In fact, the pair-wise velocity PDF of halo pairs is a complex function with skewness and kurtosis that vary substantially with scale. Using a series of collisionless N-body simulations, [107] demonstrated that the shape of the velocity PDF is determined primarily by the distribution of local densities around a halo pair, and at fixed density the velocity PDF is close to Gaussian and nearly independent of halo mass. Following this approach, we used the MDPL2 simulation [4] to calculate the over-density δ inside spheres of radius equal to the separation of the halo pairs. As δ increases, both the mean infall velocity and dispersion increase. At a fixed δ, the local velocity PDF does not show strong skewness and kurtosis, so that we can model it using a Gaussian approximation as shown in the bottom panel of figure 3.10. Therefore, one way of improving RSD models that resort to the streaming approach would be to use the N-body simulations to find the relevant quantities to be injected in the description of the pairwise PDF, such as local environment parameters. Ultimately, it could give access to the small scales of the correlation function where the effect of RSD is the largest and therefore put stronger constraints on modified

3. see table 3.1
4. see table 3.1

gravity models. [108] explored the impact of adding skewness and kurtosis in the expression of the pair-wise velocity PDF and showed that the Gaussian assumption in the streaming model breaks down on scales below $10\,h^{-1}$Mpc.

Bibliographie

[1] F. Bernardeau, S. Colombi, E. Gaztañaga, et R. Scoccimarro, *Large-scale structure of the Universe and cosmological perturbation theory*, **367**, 1 (2002), doi:10.1016/S0370-1573(02)00135-7, astro-ph/0112551.

[2] G. Lemaître, *Expansion of the universe, A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulae*, Monthly Notices of the Royal Astronomical Society **91**, 483 (1931).

[3] P. J. E. Peebles, *The large-scale structure of the universe* (1980).

[4] J. Silk, *Cosmic Black-Body Radiation and Galaxy Formation*, **151**, 459 (1968), doi:10.1086/149449.

[5] J. R. Bond et A. S. Szalay, *The collisionless damping of density fluctuations in an expanding universe*, **274**, 443 (1983), doi:10.1086/161460.

[6] E. W. Kolb et M. S. Turner, *Book-Review - the Early Universe*, **80**, 381 (1990).

[7] C. Pichon et F. Bernardeau, *Vorticity generation in large-scale structure caustics*, **343**, 663 (1999), astro-ph/9902142.

[8] O. Lahav, P. B. Lilje, J. R. Primack, et M. J. Rees, *Dynamical effects of the cosmological constant*, **251**, 128 (1991), doi:10.1093/mnras/251.1.128.

[9] E. V. Linder et R. N. Cahn, *Parameterized beyond-Einstein growth*, Astroparticle Physics **28**, 481 (2007), doi:10.1016/j.astropartphys.2007.09.003, astro-ph/0701317.

[10] Y. Gong, *Growth factor parametrization and modified gravity*, **78**, 123010 (2008), doi:10.1103/PhysRevD.78.123010, arXiv:0808.1316.

[11] Y. B. Zel'dovich, *Gravitational instability : An approximate theory for large density perturbations.*, **5**, 84 (1970).

[12] M. Crocce et R. Scoccimarro, *Renormalized cosmological perturbation theory*, **73**, 063519 (2006), doi:10.1103/PhysRevD.73.063519, astro-ph/0509418.

[13] M. Crocce et R. Scoccimarro, *Nonlinear evolution of baryon acoustic oscillations*, **77**, 023533 (2008), doi:10.1103/PhysRevD.77.023533, arXiv:0704.2783.

[14] M. Crocce, R. Scoccimarro, et F. Bernardeau, *MPTBREEZE : a fast renormalized perturbative scheme*, **427**, 2537 (2012), doi:10.1111/j.1365-2966.2012.22127.x, arXiv:1207.1465.

[15] A. Taruya, F. Bernardeau, T. Nishimichi, et S. Codis, *Direct and fast calculation of regularized cosmological power spectrum at two-loop order*, **86**, 103528 (2012), doi:10.1103/PhysRevD.86.103528, arXiv:1208.1191 [astro-ph.CO].

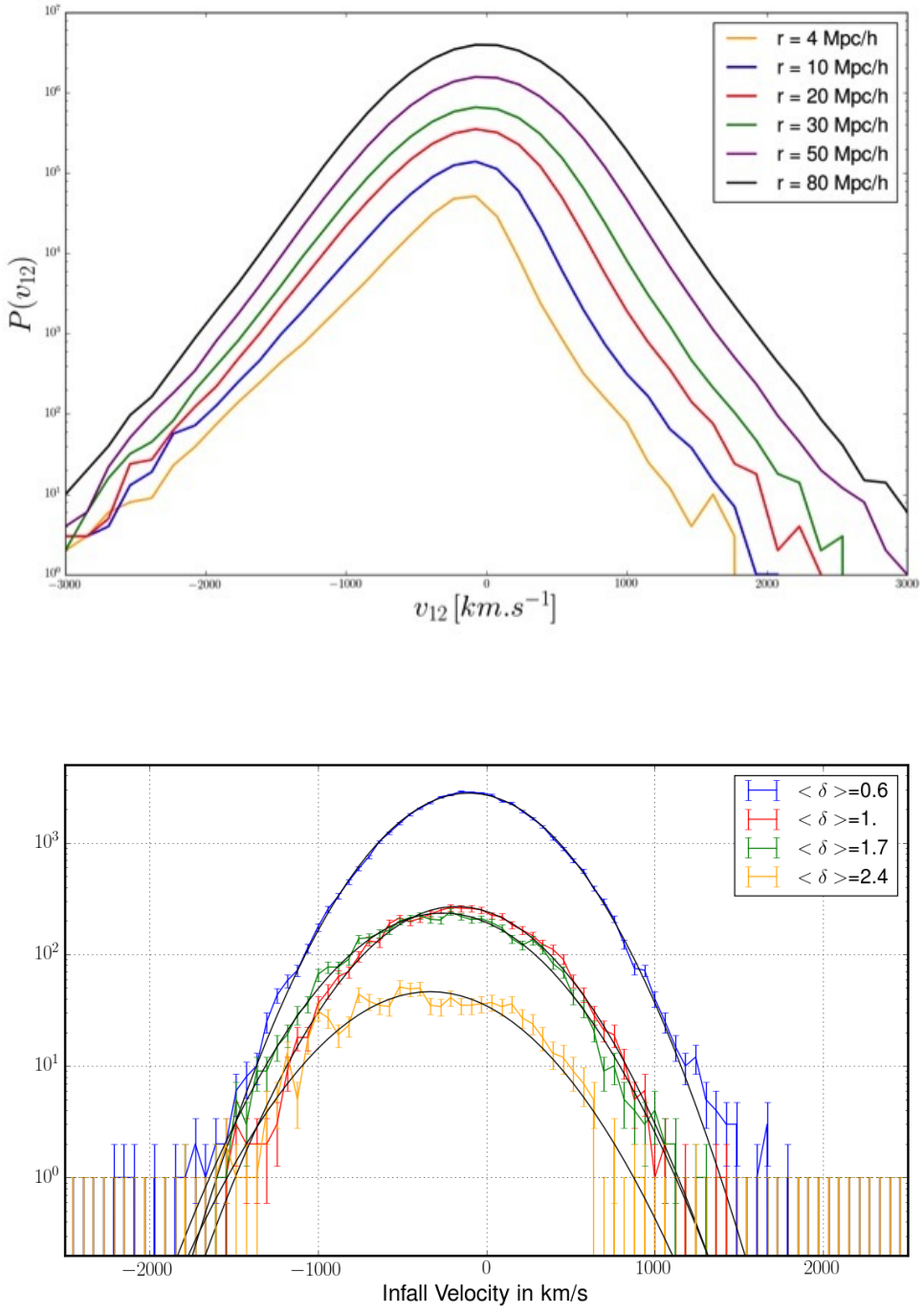


FIGURE 3.10 – Top panel : Pair-wise velocity probability distribution function at $z = 1.445$ for different pair separations using the BigMDPL simulation. Bottom panel : Density-dependent pair-wise velocity PDF using the BigMDPL simulation. Gaussianity is recovered when looking at the PDF for fixed δ.

[16] J. Carlson, M. White, et N. Padmanabhan, *Critical look at cosmological perturbation theory techniques*, **80**, 043531 (2009), doi:10.1103/PhysRevD.80.043531, arXiv:0905.0479 [astro-ph.CO].

[17] T. Matsubara, *Resumming cosmological perturbations via the Lagrangian picture : One-loop results in real space and in redshift space*, **77**, 063530 (2008), doi:10.1103/PhysRevD.77.063530, arXiv:0711.2521.

[18] H. Gil-Marín, C. Wagner, L. Verde, C. Porciani, *et al.*, *Perturbation theory approach for the power spectrum : from dark matter in real space to massive haloes in redshift space*, **11**, 029 (2012), doi:10.1088/1475-7516/2012/11/029, arXiv:1209.3771 [astro-ph.CO].

[19] J. A. Peacock et R. E. Smith, *Halo occupation numbers and galaxy bias*, **318**, 1144 (2000), doi:10.1046/j.1365-8711.2000.03779.x, astro-ph/0005010.

[20] U. Seljak, *Analytic model for galaxy and dark matter clustering*, **318**, 203 (2000), doi:10.1046/j.1365-8711.2000.03715.x, astro-ph/0001493.

[21] A. Cooray et R. Sheth, *Halo models of large scale structure*, **372**, 1 (2002), doi:10.1016/S0370-1573(02)00276-4, astro-ph/0206508.

[22] R. W. Hockney et J. W. Eastwood, *Computer simulation using particles* (1988).

[23] E. Bertschinger, *Simulations of Structure Formation in the Universe*, **36**, 599 (1998), doi:10.1146/annurev.astro.36.1.599.

[24] S. Colombi, *Dynamics of the large-scale structure of the universe : /N-body techniques*, **45**, 373 (2001), doi:10.1016/S1387-6473(00)00159-7.

[25] W. Dehnen, *Towards optimal softening in three-dimensional N-body codes - I. Minimizing the force error*, **324**, 273 (2001), doi:10.1046/j.1365-8711.2001.04237.x, astro-ph/0011568.

[26] A. Klypin et J. Holtzman, *Particle-Mesh code for cosmological simulations*, ArXiv Astrophysics e-prints (1997), astro-ph/9712217.

[27] K. Riebe, A. M. Partl, H. Enke, J. Forero-Romero, *et al.*, *The MultiDark Database : Release of the Bolshoi and MultiDark cosmological simulations*, Astronomische Nachrichten **334**, 691 (2013), doi:10.1002/asna.201211900.

[28] M. Davis, G. Efstathiou, C. S. Frenk, et S. D. M. White, *The evolution of large-scale structure in a universe dominated by cold dark matter*, **292**, 371 (1985), doi:10.1086/163168.

[29] A. Knebe, S. R. Knollmann, S. I. Muldrew, F. R. Pearce, *et al.*, *Haloes gone MAD : The Halo-Finder Comparison Project*, **415**, 2293 (2011), doi:10.1111/j.1365-2966.2011.18858.x, arXiv:1104.0949.

[30] P. S. Behroozi, R. H. Wechsler, et H.-Y. Wu, *The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores*, **762**, 109 (2013), doi:10.1088/0004-637X/762/2/109, arXiv:1110.4372 [astro-ph.CO].

[31] M. White, A. D. Myers, N. P. Ross, D. J. Schlegel, *et al.*, *The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey*, **424**, 933 (2012), doi:10.1111/j.1365-2966.2012.21251.x, arXiv:1203.5306.

[32] S. Eftekharzadeh, A. D. Myers, M. White, D. H. Weinberg, *et al.*, *Clustering of intermediate redshift quasars using the final SDSS III-BOSS sample*, ArXiv e-prints (2015), arXiv:1507.08380.

[33] P. Laurent, S. Eftekharzadeh, J.-M. Le Goff, A. Myers, *et al.*, *Clustering of quasars in SDSS-IV eBOSS : study of potential systematics and bias determination*, **7**, 017 (2017), doi:10.1088/1475-7516/2017/07/017, arXiv:1705.04718.

[34] S. A. Rodríguez-Torres, J. Comparat, F. Prada, G. Yepes, *et al.*, *Clustering of quasars in the first year of the SDSS-IV eBOSS survey : interpretation and halo occupation distribution*, **468**, 728 (2017), doi:10.1093/mnras/stx454, arXiv:1612.06918.

[35] A. Klypin, G. Yepes, S. Gottlöber, F. Prada, *et al.*, *MultiDark simulations : the story of dark matter halo concentrations and density profiles*, **457**, 4340 (2016), doi:10.1093/mnras/stw248, arXiv:1411.4001.

[36] V. Springel, *The cosmological simulation code GADGET-2*, **364**, 1105 (2005), doi:10.1111/j.1365-2966.2005.09655.x, astro-ph/0505010.

[37] J. Comparat, F. Prada, G. Yepes, et A. Klypin, *Accurate mass and velocity functions of dark matter haloes*, **469**, 4157 (2017), doi:10.1093/mnras/stx1183, arXiv:1702.01628.

[38] S. Habib *et al.*, *HACC : Simulating Sky Surveys on State-of-the-Art Supercomputing Architectures*, New Astron. **42**, 49 (2016), doi:10.1016/j.newast.2015.06.003, arXiv:1410.2805 [astro-ph.IM].

[39] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, *et al.*, *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Cosmological Interpretation*, **192**, 18 (2011), doi:10.1088/0067-0049/192/2/18, arXiv:1001.4538 [astro-ph.CO].

[40] J. E. Gunn et J. R. Gott, III, *On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution*, **176**, 1 (1972), doi:10.1086/151605.

[41] C. Lacey et S. Cole, *Merger rates in hierarchical models of galaxy formation*, **262**, 627 (1993), doi:10.1093/mnras/262.3.627.

[42] G. L. Bryan et M. L. Norman, *Statistical Properties of X-Ray Clusters : Analytic and Numerical Comparisons*, **495**, 80 (1998), doi:10.1086/305262, astro-ph/9710107.

[43] J. M. Bardeen, J. R. Bond, N. Kaiser, et A. S. Szalay, *The statistics of peaks of Gaussian random fields*, **304**, 15 (1986), doi:10.1086/164143.

[44] J. F. Navarro, C. S. Frenk, et S. D. M. White, *The Structure of Cold Dark Matter Halos*, **462**, 563 (1996), doi:10.1086/177173, astro-ph/9508025.

[45] L. Hernquist, *An analytical model for spherical galaxies and bulges*, **356**, 359 (1990), doi:10.1086/168845.

[46] J. F. Navarro, E. Hayashi, C. Power, A. R. Jenkins, *et al.*, *The inner structure of ΛCDM haloes - III. Universality and asymptotic slopes*, **349**, 1039 (2004), doi:10.1111/j.1365-2966.2004.07586.x, astro-ph/0311231.

[47] L. Gao, J. F. Navarro, S. Cole, C. S. Frenk, *et al.*, *The redshift dependence of the structure of massive Λ cold dark matter haloes*, **387**, 536 (2008), doi:10.1111/j.1365-2966.2008.13277.x, arXiv:0711.0746.

[48] W. H. Press et P. Schechter, *Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation*, **187**, 425 (1974), doi:10.1086/152650.

[49] J. R. Bond, S. Cole, G. Efstathiou, et N. Kaiser, *Excursion set mass functions for hierarchical Gaussian fluctuations*, **379**, 440 (1991), doi:10.1086/170520.

[50] R. K. Sheth et G. Tormen, *Large-scale bias and the peak background split*, **308**, 119 (1999), doi:10.1046/j.1365-8711.1999.02692.x, astro-ph/9901122.

[51] R. K. Sheth, H. J. Mo, et G. Tormen, *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes*, **323**, 1 (2001), doi:10.1046/j.1365-8711.2001.04006.x, astro-ph/9907024.

[52] J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, *et al.*, *Toward a Halo Mass Function for Precision Cosmology : The Limits of Universality*, **688**, 709 (2008), doi:10.1086/591439, arXiv:0803.2706.

[53] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, *et al.*, *Stable clustering, the halo model and non-linear cosmological power spectra*, **341**, 1311 (2003), doi:10.1046/j.1365-8711.2003.06503.x, astro-ph/0207664.

[54] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, *et al.*, *Revising the Halofit Model for the Nonlinear Matter Power Spectrum*, **761**, 152 (2012), doi:10.1088/0004-637X/761/2/152, arXiv:1208.2701.

[55] J. N. Fry et E. Gaztanaga, *Biasing and hierarchical statistics in large-scale structure*, **413**, 447 (1993), doi:10.1086/173015, astro-ph/9302009.

[56] P. McDonald et A. Roy, *Clustering of dark matter tracers : generalizing bias for the coming era of precision LSS*, **8**, 020 (2009), doi:10.1088/1475-7516/2009/08/020, arXiv:0902.0991 [astro-ph.CO].

[57] H. Gil-Marín, J. Guy, P. Zarrouk, E. Burtin, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : structure growth rate measurement from the anisotropic quasar power spectrum in the redshift range 0.8 z 2.2*, **477**, 1604 (2018), doi:10.1093/mnras/sty453, arXiv:1801.02689.

[58] T. Matsubara, *Nonlinear perturbation theory with halo bias and redshift-space distortions via the Lagrangian picture*, **78**, 083519 (2008), doi:10.1103/PhysRevD.78.083519, arXiv:0807.1733.

[59] R. H. Wechsler, M. A. K. Gross, J. R. Primack, G. R. Blumenthal, *et al.*, *Implications of Spikes in the Redshift Distribution of Z ˜ 3 Galaxies*, **506**, 19 (1998), doi:10.1086/306229, astro-ph/9712141.

[60] P. Colín, A. A. Klypin, A. V. Kravtsov, et A. M. Khokhlov, *Evolution of Bias in Different Cosmological Models*, **523**, 32 (1999), doi:10.1086/307710, astro-ph/9809202.

[61] A. V. Kravtsov, A. A. Berlind, R. H. Wechsler, A. A. Klypin, *et al.*, *The Dark Side of the Halo Occupation Distribution*, **609**, 35 (2004), doi:10.1086/420959, astro-ph/0308519.

[62] A. Vale et J. P. Ostriker, *Linking halo mass to galaxy luminosity*, **353**, 189 (2004), doi:10.1111/j.1365-2966.2004.08059.x, astro-ph/0402500.

[63] S. Trujillo-Gomez, A. Klypin, J. Primack, et A. J. Romanowsky, *Galaxies in ΛCDM with Halo Abundance Matching : Luminosity-Velocity Relation, Baryonic Mass-Velocity Relation, Velocity Function, and Clustering*, **742**, 16 (2011), doi:10.1088/0004-637X/742/1/16, arXiv:1005.1289.

[64] G. Favole, J. Comparat, F. Prada, G. Yepes, *et al.*, *Clustering properties of g-selected galaxies at z*
0.8, **461**, 3421 (2016), doi:10.1093/mnras/stw1483, arXiv:1507.04356.

[65] S. E. Nuza, A. G. Sánchez, F. Prada, A. Klypin, *et al.*, *The clustering of galaxies at z 0.5 in the SDSS-III Data Release 9 BOSS-CMASS sample : a test for the ΛCDM cosmology*, **432**, 743 (2013), doi:10.1093/mnras/stt513, arXiv:1202.6057.

[66] A. A. Berlind et D. H. Weinberg, *The Halo Occupation Distribution : Toward an Empirical Determination of the Relation between Galaxies and Mass*, **575**, 587 (2002), doi:10.1086/341469, astro-ph/0109001.

[67] Y. Shen, M. A. Strauss, M. Oguri, J. F. Hennawi, *et al.*, *Clustering of High-Redshift (z = 2.9) Quasars from the Sloan Digital Sky Survey*, **133**, 2222 (2007), doi:10.1086/513517, astro-ph/0702214.

[68] J. Richardson, Z. Zheng, S. Chatterjee, D. Nagai, *et al.*, *The Halo Occupation Distribution of SDSS Quasars*, **755**, 30 (2012), doi:10.1088/0004-637X/755/1/30, arXiv:1203.4570.

[69] J. L. Tinker, E. S. Sheldon, R. H. Wechsler, M. R. Becker, *et al.*, *Cosmological Constraints from Galaxy Clustering and the Mass-to-number Ratio of Galaxy Clusters*, **745**, 16 (2012), doi:10.1088/0004-637X/745/1/16, arXiv:1104.1635.

[70] N. P. Ross, Y. Shen, M. A. Strauss, D. E. Vanden Berk, *et al.*, *Clustering of Low-redshift (z = 2.2) Quasars from the Sloan Digital Sky Survey*, **697**, 1634 (2009), doi:10.1088/0004-637X/697/2/1634, arXiv:0903.3230.

[71] I. Kayo et M. Oguri, *Very small scale clustering of quasars from a complete quasar lens survey*, **424**, 1363 (2012), doi:10.1111/j.1365-2966.2012.21321.x, arXiv:1203.6410.

[72] D. B. Sanders, B. T. Soifer, J. H. Elias, B. F. Madore, *et al.*, *Ultraluminous infrared galaxies and the origin of quasars*, **325**, 74 (1988), doi:10.1086/165983.

[73] R. G. Carlberg, *Quasar evolution via galaxy mergers*, **350**, 505 (1990), doi:10.1086/168406.

[74] P. F. Hopkins, R. S. Somerville, L. Hernquist, T. J. Cox, *et al.*, *The Relation between Quasar and Merging Galaxy Luminosity Functions and the Merger-driven Star Formation History of the Universe*, **652**, 864 (2006), doi:10.1086/508503, astro-ph/0602290.

[75] R. H. Wechsler et J. L. Tinker, *The Connection between Galaxies and their Dark Matter Halos*, ArXiv e-prints (2018), arXiv:1804.03097.

[76] S. de la Torre et L. Guzzo, *Modelling non-linear redshift-space distortions in the galaxy clustering pattern : systematic errors on the growth rate parameter*, **427**, 327 (2012), doi:10.1111/j.1365-2966.2012.21824.x, arXiv:1202.5559.

[77] A. P. Hearin, A. R. Zentner, F. C. van den Bosch, D. Campbell, *et al.*, *Introducing decorated HODs : modelling assembly bias in the galaxy-halo connection*, **460**, 2552 (2016), doi:10.1093/mnras/stw840, arXiv:1512.03050.

[78] B. Diemer, S. More, et A. V. Kravtsov, *The Pseudo-evolution of Halo Mass*, **766**, 25 (2013), doi:10.1088/0004-637X/766/1/25, arXiv:1207.0816 [astro-ph.CO].

[79] B. Diemer et A. V. Kravtsov, *Dependence of the Outer Density Profiles of Halos on Their Mass Accretion Rate*, **789**, 1 (2014), doi:10.1088/0004-637X/789/1/1, arXiv:1401.1216.

[80] P. Mansfield, A. V. Kravtsov, et B. Diemer, *Splashback Shells of Cold Dark Matter Halos*, **841**, 34 (2017), doi:10.3847/1538-4357/aa7047, arXiv:1612.01531.

[81] N. Kaiser, *Clustering in real space and in redshift space*, **227**, 1 (1987).

[82] R. Scoccimarro, *Redshift-space distortions, pairwise velocities, and nonlinearities*, **70**, 083007 (2004), doi:10.1103/PhysRevD.70.083007, astro-ph/0407214.

[83] J. A. Peacock et M. J. West, *The power spectrum of Abell cluster correlations*, **259**, 494 (1992), doi:10.1093/mnras/259.3.494.

[84] A. Taruya, T. Nishimichi, et S. Saito, *Baryon acoustic oscillations in 2D : Modeling redshift-space power spectrum from perturbation theory*, **82**, 063522 (2010), doi:10.1103/PhysRevD.82.063522, arXiv:1006.0699.

[85] K. C. Chan et R. Scoccimarro, *Halo sampling, local bias, and loop corrections*, **86**, 103519 (2012), doi:10.1103/PhysRevD.86.103519, arXiv:1204.5770 [astro-ph.CO].

[86] K. B. Fisher, M. Davis, M. A. Strauss, A. Yahil, *et al.*, *Clustering in the 1.2-Jy IRAS galaxy redshift survey - II. Redshift distortions and dzeta(rp, pi).*, **267**, 927 (1994), doi:10.1093/mnras/267.4.927.

[87] K. B. Fisher, *On the Validity of the Streaming Model for the Redshift-Space Correlation Function in the Linear Regime*, **448**, 494 (1995), doi:10.1086/175980, astro-ph/9412081.

[88] B. A. Reid et M. White, *Towards an accurate model of the redshift-space clustering of haloes in the quasi-linear regime*, **417**, 1913 (2011), doi:10.1111/j.1365-2966.2011.19379.x, arXiv:1105.4165.

[89] B. A. Reid, L. Samushia, M. White, W. J. Percival, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : measurements of the growth of structure and expansion rate at $z = 0.57$ from anisotropic clustering*, **426**, 2719 (2012), doi:10.1111/j.1365-2966.2012.21779.x, arXiv:1203.6641.

[90] J. Carlson, B. Reid, et M. White, *Convolution Lagrangian perturbation theory for biased tracers*, **429**, 1674 (2013), doi:10.1093/mnras/sts457, arXiv:1209.0780.

[91] L. Wang, B. Reid, et M. White, *An analytic model for redshift-space distortions*, **437**, 588 (2014), doi:10.1093/mnras/stt1916, arXiv:1306.1804.

[92] S. Cole et N. Kaiser, *Biased clustering in the cold dark matter cosmogony*, **237**, 1127 (1989), doi:10.1093/mnras/237.4.1127.

[93] M. White, B. Reid, C.-H. Chuang, J. L. Tinker, *et al.*, *Tests of redshift-space distortions models in configuration space for the analysis of the BOSS final data release*, **447**, 234 (2015), doi:10.1093/mnras/stu2460, arXiv:1408.5435.

[94] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, *et al.*, *The power spectrum and bispectrum of SDSS DR11 BOSS galaxies - I. Bias and gravity*, **451**, 539 (2015), doi:10.1093/mnras/stv961, arXiv:1407.5668.

[95] H. Gil-Marín, W. J. Percival, J. R. Brownstein, C.-H. Chuang, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : RSD measurement from the LOS-dependent power spectrum of DR12 BOSS galaxies*, **460**, 4188 (2016), doi:10.1093/mnras/stw1096, arXiv:1509.06386.

[96] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies*, **465**, 1757 (2017), doi:10.1093/mnras/stw2679, arXiv:1606.00439.

[97] F. Beutler, S. Saito, H.-J. Seo, J. Brinkmann, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : testing gravity with redshift space distortions using the power spectrum multipoles*, **443**, 1065 (2014), doi:10.1093/mnras/stu1051, arXiv:1312.4611.

[98] F. Beutler, H.-J. Seo, S. Saito, C.-H. Chuang, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : anisotropic galaxy clustering in Fourier space*, **466**, 2242 (2017), doi:10.1093/mnras/stw3298, arXiv:1607.03150.

[99] A. G. Sánchez, R. Scoccimarro, M. Crocce, J. N. Grieb, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : Cosmological implications of the configuration-space clustering wedges*, **464**, 1640 (2017), doi:10.1093/mnras/stw2443, arXiv:1607.03147.

[100] J. N. Grieb, A. G. Sánchez, S. Salazar-Albornoz, R. Scoccimarro, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : Cosmological implications of the Fourier space wedges of the final sample*, **467**, 2085 (2017), doi:10.1093/mnras/stw3384, arXiv:1607.03143.

[101] J. Hou, A. G. Sánchez, R. Scoccimarro, S. Salazar-Albornoz, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : anisotropic clustering analysis in configuration-space*, ArXiv e-prints (2018), arXiv:1801.02656.

[102] L. Samushia, B. A. Reid, M. White, W. J. Percival, *et al.*, *The clustering of galaxies in the SDSS-III DR9 Baryon Oscillation Spectroscopic Survey : testing deviations from Λ and general relativity using anisotropic clustering of galaxies*, **429**, 1514 (2013), doi:10.1093/mnras/sts443, arXiv:1206.5309.

[103] L. Samushia, B. A. Reid, M. White, W. J. Percival, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : measuring growth rate and geometry with anisotropic clustering*, **439**, 3504 (2014), doi:10.1093/mnras/stu197, arXiv:1312.4899.

[104] S. Alam, F. D. Albareti, C. Allende Prieto, F. Anders, *et al.*, *The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey : Final Data from SDSS-III*, ArXiv e-prints (2015), arXiv:1501.00963.

[105] S. Satpathy, S. Alam, S. Ho, M. White, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : on the measurement of growth rate using galaxy correlation functions*, **469**, 1369 (2017), doi:10.1093/mnras/stx883, arXiv:1607.03148.

[106] P. Zarrouk, E. Burtin, H. Gil-Marín, A. J. Ross, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2*, **477**, 1639 (2018), doi:10.1093/mnras/sty506.

[107] J. L. Tinker, *Redshift-space distortions with the halo occupation distribution - II. Analytic model*, **374**, 477 (2007), doi:10.1111/j.1365-2966.2006.11157.x, astro-ph/0604217.

[108] D. Bianchi, W. J. Percival, et J. Bel, *Improving the modelling of redshift-space distortions- II. A pairwise velocity model covering large and small scales*, **463**, 3783 (2016), doi:10.1093/mnras/stw2243, arXiv:1602.02780.

Chapitre 4

Study of potential systematics

Contents

Identification and elimination of systematic uncertainties are essential to reach the best-to-date precision on the cosmological measurements from clustering analysis. The purpose of this chapter is to present and minimize the impact of sources of systematic uncertainties on the anisotropic clustering of the eBOSS quasar sample, in order to provide robust constraints on the cosmological parameters. It is usually reached when the effect of systematic uncertainties represents less than 20% of the statistical precision.

In order to investigate the different origins of systematic effects, we generated different sets of mock catalogs that are fictional realizations of the data used as a benchmark

to test the pipeline. Section 4.1 explains the methodology we adopt by specifying the RSD model used to model the redshift-space correlation function, the fitting conditions and the covariance matrix used to infer the uncertainty on the fitting parameters which is obtained from approximate mock catalogs. Then, we divide the systematics into two types : modelling systematics in section 4.2 - which include the systematic uncertainty associated with the RSD model, the bias model and the impact of redshift uncertainties - and observational systematics in section 4.3 that gather systematics related to the imaging conditions, the observational strategy and instrumental limitations. This chapter is based on the analysis performed in [1].

4.1 Methodology

4.1.1 Configuration space multipoles and wedges

The two-point correlation function $\xi(s, \mu)$ is calculated using the Landy-Szalay estimator defined in equation 6.6 for the two variables s and μ. The public code CUTE [2] is used to calculate paircounts as a function of comoving separation s, and μ from 0 to 200 h^{-1}Mpc subdivided into 25 bins which makes a cell size of 8 h^{-1}Mpc and 30 bins in μ from 0 to 1.

The two-point correlation function is projected onto the Legendre polynomial basis [3] through :

$$\xi_l(s) = \frac{2l+1}{2} \sum_j \xi(s, \mu_j) P_l(\mu_j) d\mu \,, \tag{4.1}$$

where only $l = 0, 2, 4$ are the monopole, quadrupole and hexadecapole respectively.. The analysis can also be performed by cutting the domain in μ into “wedges" [4] :

$$\xi_{wi}(s) = \frac{1}{\mu_{i,\max} - \mu_{i,\min}} \int_{\mu_{i,\min}}^{\mu_{i,\max}} \xi(s, \mu) d\mu \,. \tag{4.2}$$

In the present work, we consider three wedges in μ of constant size $\Delta\mu = 1/3$ where the wedge $2/3 < \mu < 1$ corresponds to the closest region to the LOS.

The theoretical predictions are computed using the CLPT-GS model presented in section 3.4.2. To summarize the differents steps :

(i) The linear power spectrum is obtained using CAMB with the fiducial cosmology given in table 4.1

(ii) We use the CLPT code [1] to calculate $\xi(r)$, v_{12}, σ_{12} according to equation 6.12. The CLPT code takes a linear power spectrum as input and provides a file for each real-space observable.

(iii) These quantities are then used in our PYTHON code that performs the integral given by the Gaussian streaming model in equation 6.10. We then compute the theoretical predictions for the Legendre multipoles and wedges of the two-point correlation function in redshif space.

The fits are thus performed with the CLPT-GS model, under the same conditions for the data and the mock catalogs from 16 h^{-1}Mpc to 136 h^{-1}Mpc using binwidth of 8 h^{-1}Mpc. The priors on the fit parameters are :

1. https://github.com/wll745881210/CLPTGSRSD.git

parameter	prior
F'	flat prior, range (0.1, 2.8)
f	flat prior, range (0,5)
$\alpha_{\parallel}$	flat prior, range (0,2)
$\alpha_{\perp}$	flat prior, range (0,2)
$\sigma_{\rm tot}$	flat prior, range (0,20)

4.1.2 The covariance matrix

Inference of cosmological information from large-scale structure observations requires the estimation of the covariance of the measurements as accurately as possible. The covariance matrix is estimated from a set of N_m mock measurements, denoted $\xi^{(n)}$, where $n \in 1, ..., N_m$, so that :

$$C_{A,B} = \frac{1}{N_m - 1} \sum_{n=1}^{N_m} [\xi_A^{(n)}(s_i) - \bar{\xi}_A(s_i)] \, [\xi_B^{(n)}(s_j) - \bar{\xi}_B(s_j)] \tag{4.3}$$

where N_m is the number of mock realizations with 15 bins per multipole/wedge, so the dimensions of the covariance matrix for the NGC+SGC is (3×15)×(3×15).
The mean of the correlation function is defined by :

$$\bar{\xi}_A(s_i) = \frac{1}{N_m} \sum_{n=1}^{N_m} \xi_A^n(s_i) \tag{4.4}$$

and the correlation coefficients of the correlation matrix are :

$$r_{A,B} = \frac{C_{A,B}}{\sqrt{C_{AA} C_{BB}}} \tag{4.5}$$

In this analysis, two sets of synthetic mock catalogs that mimic the clustering properties of the eBOSS quasar sample are used to obtain the covariance matrix of the measurements. To do so, these catalogs incorporate the full survey geometry (i.e., they have the same angular and radial selection function as the data) and important observational systematics such as the veto masks and fibre collisions. A precise estimate of the covariance matrix requires a large number of mock catalogs, so we need to generate about hundreds to thousands of realizations. Contrary to N-body simulations, those mock catalogs can rely on fast approximate gravity solvers to generate halo catalogs because their purpose is not to test and model accurately the non-linear evolution of LSS formation. These halos are then populated with synthetic galaxies using the empirical models described in section 3.3.3 whose parameters are tuned to match the clustering of data.

Table 4.1 summarizes the cosmological parameters used to generate the initial density field in each set of mock catalogs. QPM and EZ refer to two techniques used to construct fast approximate mock catalogs, Outer Rim is the N-body simulations we use to build accurate mock catalogs.

4.1.2.1 QPM mock catalogs

A first set of approximate catalogs relies on a technique known as quick-particle-mesh [QPM 5]. This method uses a low-resolution particle mesh gravity solver to generate the large-scale dark matter density field from initial conditions that have been created using the cosmological parameters given as 'QPM' in table 4.1 and then evolve it using

TABLE 4.1 – Cosmological parameters of the fiducial cosmological model compared to the one of N-body simulations used in the analysis.

Configuration	Ω_m	h	$\Omega_b h^2$	σ_8
Fiducial	0.31	0.676	0.022	0.8
QPM	0.29	0.7	0.02247	0.8
EZ	0.307115	0.6777	0.02214	0.8228
Outer Rim	0.26479	0.71	0.02258	0.8

2LPT. This approach does not have sufficient spatial and mass resolution to form dark matter halos so particles are sampled from the density field in a way that mimics both one-point and two-point distribution of halos, in other words the mass function and bias function. Then, in a second step, these halos are populated with quasars according to the five-parameter HOD [6] described in section 3.3.3 with a sixth parameter that accounts for the duty cycle of quasars :

$$\langle N_{\rm cen}\rangle_M = \tau \cdot \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm cen}}{\log \sigma_M}\right)\right] \tag{4.6}$$

$$\langle N_{\rm sat}\rangle_M = \left(\frac{M}{M_{\rm sat}}\right)^{\alpha_{\rm sat}} \cdot \exp\left(-\frac{M_{\rm cut}}{M}\right) \tag{4.7}$$

where $\langle N_{\rm cen}\rangle_M$ is the probability for a halo of mass M to host a central quasar and $\langle N_{\rm sat}\rangle_M$ is the number of satellite in a halo of mass M. The values of the parameters that reproduce the data are thus $M_{\rm cen} = 1.35\,10^{12}\,M_\odot$, $\log \sigma_M = 0.2$, $\alpha_{\rm sat} = 1$, $M_{\rm cut} = 10^8\,M_\odot$ and $M_{\rm sat} = 1.93 \times 10^{15}\,M_\odot$, which results in a population that consists of $\simeq 13\%$ satellites. Under these conditions the typical mass for dark matter halos hosting quasars is $M_{\rm cen} = 10^{12.5}\,M_\odot$.

One hundred realizations have been generated using cubic boxes of size $L = 5120\,h^{-1}$.Mpc which produced catalogs with positions (x,y,z) and velocities (v_x,v_y,v_z). The same 100 cubic boxes are used for the NGC and SGC. Then, these cubic boxes have then be remapped to be parallelepipoid in order to fit the volume of the full planned survey using the remapping technique implemented in the code MAKE SURVEY [7, 5].

Figure 4.6 is an illustration of the remapping technique that transforms a cubic box into a parallepipoid by moving sub-regions of the cube to another place. [7] showed that this remapping techniques preserved the structures without adding correlations between regions that were not connected. It is based on the observation that a cube with periodic boundary conditions is equivalent to an infinite 3D space with discrete translational symmetry.

The analysis presented in this work is based on the first two years of data taking of eBOSS so we use a smaller volume which allows us to use different parts of a single box to produce more realizations. By changing the orientation of the original box and varying the different combinations to rotate it, we identified four configurations with less than 1.5% overlap. Using these 4 rotations allows us to produce 400 QPM mocks per Galactic cap. The overlap between NGC and SGC can be as high as 10% but we identified the pairs of configurations for which the overlap is less than 2% in order to combine them and to obtain the full DR14 sample. Once the parallelepipoid boxes have been obtained, the cartesian comoving coordinates (x,y,z) are transformed to angular coordinates and radial distance (RA,DEC,z) using a flat ΛCDM cosmology with parameters in table 4.1. The angular and veto masks are also applied as in the data and the number density of objects in the remaining area is downsampled to fit the redshift distribution of the data.

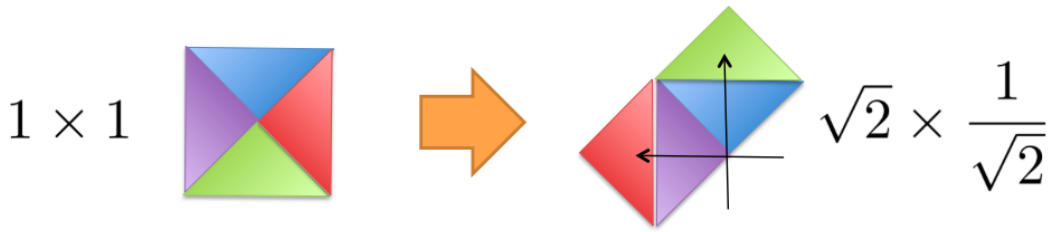


FIGURE 4.1 – Illustration of the remapping technique : after translation, a cubic box of size 1×1 is transformed into a parallepipoid of size $\sqrt{2} \times 1/\sqrt{2}$. Credit : Martin White http://mwhite.berkeley.edu/BoxRemap/.

As explained in section 2.5, quasars suffer from important redshift uncertainties that are taken into account in the QPM mocks by smearing the redshifts according to a Gaussian distribution of width taken from the eBOSS requirements [8], namely $\sigma_z = 300\,\mathrm{km.s^{-1}}$ for $z < 1.5$ and $\sigma_z = 300 + 400 \times (z - 1.5)\,\mathrm{km.s^{-1}}$ for $z > 1.5$.

4.1.2.2 EZ mock catalogs

The QPM mocks belong to the class of approximate catalogs that rely on Lagrangian methods where initial conditions are generated from a linear power spectrum and then galaxies are found deterministically using empirical prescriptions on halos. Another method consists in generating a density field and populating it with halos using a biasing scheme that can be non-linear and stochastic. This is the case for the Effective Zel'dovich mock catalogs [EZ, 9] that are constructed using the Zel'dovich approximation of the density field and an effective modelling with few parameters, which can be efficiently calibrated with observations or numerical simulations. The performance of a wide variety of mock halo and galaxy catalog generators have been compared in [9] and they demonstrated that the EZ mock technique can reproduce the clustering of a given sample (including two-point and three-point statistics) with low memory requirements compared to other methods.

For this analysis, the bias parameters are calibrated on the observed eBOSS DR14 quasar clustering directly and figure 4.2 shows the comparison between the linear bias measured from EZ mocks and other measurements [10, 11]. The linear bias b is related to the first Lagrangian parameter F' of the biasing scheme used in this work by $b = 1 + F'$. A comparison between different bias measurements of quasar samples will be provided in section 5.4.3.

The relevant parameters were tuned on the data for the NGC and SGC separately to reproduce the difference in clustering as shown in figure 4.3. The mean of the 1,000 EZ mocks is shown in dashed line and the error bars on the data points correspond to the square of the diagonal elements of the inverse of the covariance matrix. Given that the EZ mocks are adjusted on the data directly, they already contain the redshift resolution of the data that affects the clustering, especially the quadrupole on scales below $\sim 40\,h^{-1}$Mpc. We use 1000 independent realizations for each Galactic cap. One thousand light-cone mock catalogs covering the full survey area of DR14 have been produced using boxes of size $L = 5000\,h^{-1}$.Mpc and an initial power spectrum based on ΛCDM cosmology with $\Omega_\mathrm{m} = 0.307115$, $h = 0.6777$, $\sigma_8 = 0.8225$, $\Omega_\mathrm{b} = 0.048206$ and $n_s = 0.9611$. It corresponds to the cosmology used to produce a set of BOSS mock catalogs (see [12] for more details). Each light-cone is composed of seven redshift shells, each shell shares the same initial Gaussian density but has different EZ mocks parameters to account for the redshift evolution of the

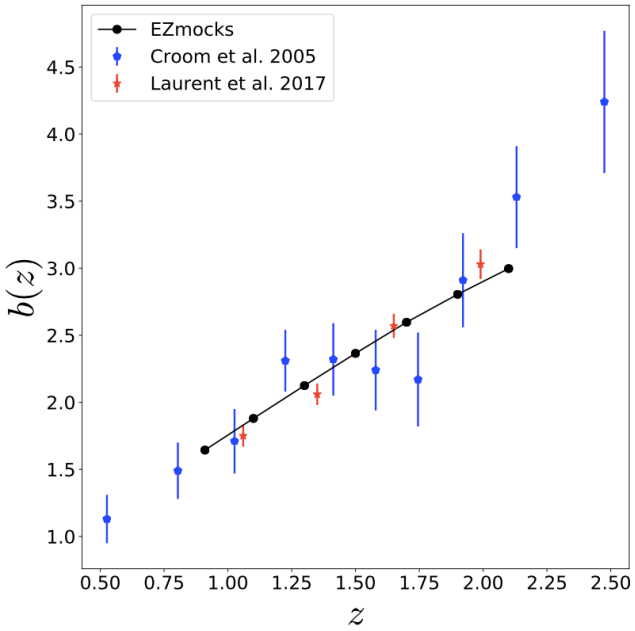


FIGURE 4.2 – Linear bias measurements from observed data using the 2dF QSO Redshift Survey [10] and the first year of eBOSS data taking [11] compared to the linear bias measured in the EZ mocks used in this analysis. The bias evolution in the EZ mocks agrees with these works.

observed quasar clustering. So in total, we generated 1000×7(shells)$\times 2$(NGC+SGC)$= 14,000$ EZ mock boxes and we used the code MAKE SURVEY to construct each redshift shell from the corresponding box. Further details on the updated algorithm, clustering performance and implementation of the redshift evolution will be provided in Chuang et al. (in prep).

We include some synthetic observational features in the EZ and QPM mocks in order to have a more realistic covariance matrices and to quantify the effects of those features on the cosmological parameters of interest. The production of the improved set of EZ mocks is explained in section 4.3.1 and the correlation matrices we infer from these synthetic mock catalogs are displayed in Figure 4.4 for the 3-multipole and 3-wedge analyses. The most important correlations are along the diagonal but the inclusion of the off-diagonal terms matters to obtain a more precise measurement of the cosmological parameters.

4.1.3 Fitting conditions

In order to obtain a better statistical precision in the cosmological parameters and to be less dominated by statistical fluctuations, we combine the NGC and SGC of the data set (either the EZ mocks or the DR14 sample) and perform a fitting on the combined NGC+SGC. As shown in figure 2.19, the two caps do not overlap and we can assume their statistics are not correlated, so we consider an area-averaged estimator for the NGC+SGC defined by :

$$\xi_{\mathrm{NGC+SGC}} = \frac{\xi_{\mathrm{NGC}} \cdot A_{\mathrm{NGC}} + \xi_{\mathrm{SGC}} \cdot A_{\mathrm{SGC}}}{A_{\mathrm{NGC}} + A_{\mathrm{SGC}}} \tag{4.8}$$

where A_{NGC} (resp. A_{SGC}) is the area covered by the NGC (resp. SGC).

The cosmological information is extracted through a fit of the measured correlation function with the CLPT-GS model which gives a prediction for the correlation function

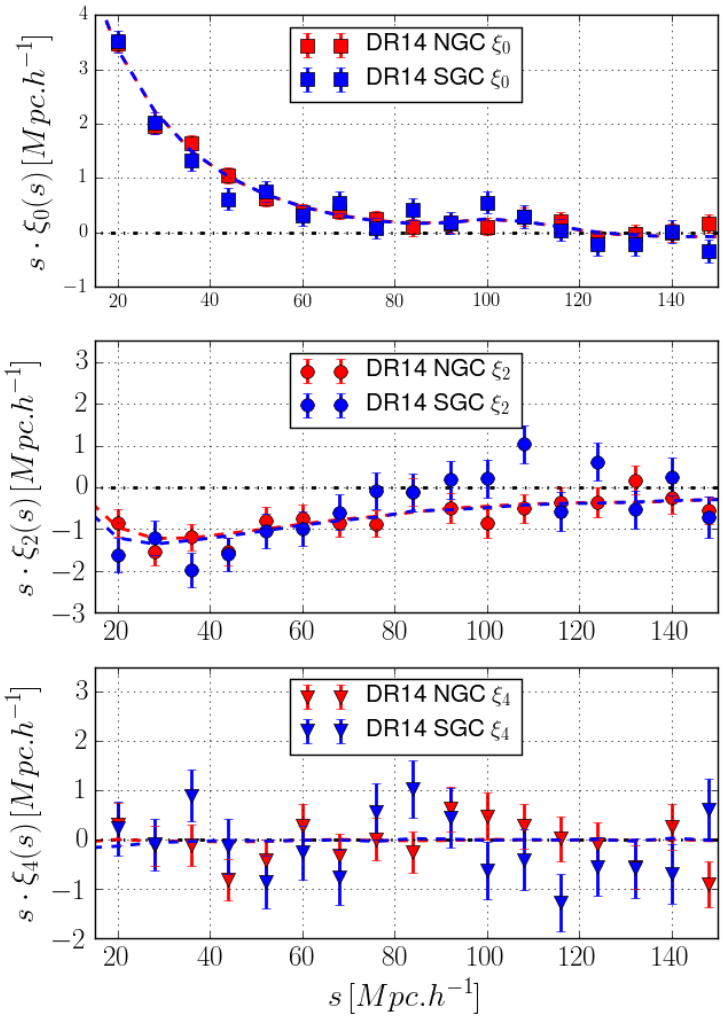


FIGURE 4.3 – Top panel : Monopole of the eBOSS DR14 NGC (blue) and SGC (red) compared to the mean of the 1000 EZ mocks (dashed). Middle and bottom panels : Same for the quadrupole and the hexadecapole. The parameters of the EZ mocks are tuned on the observed clustering of the data for each Galactic cap separately.

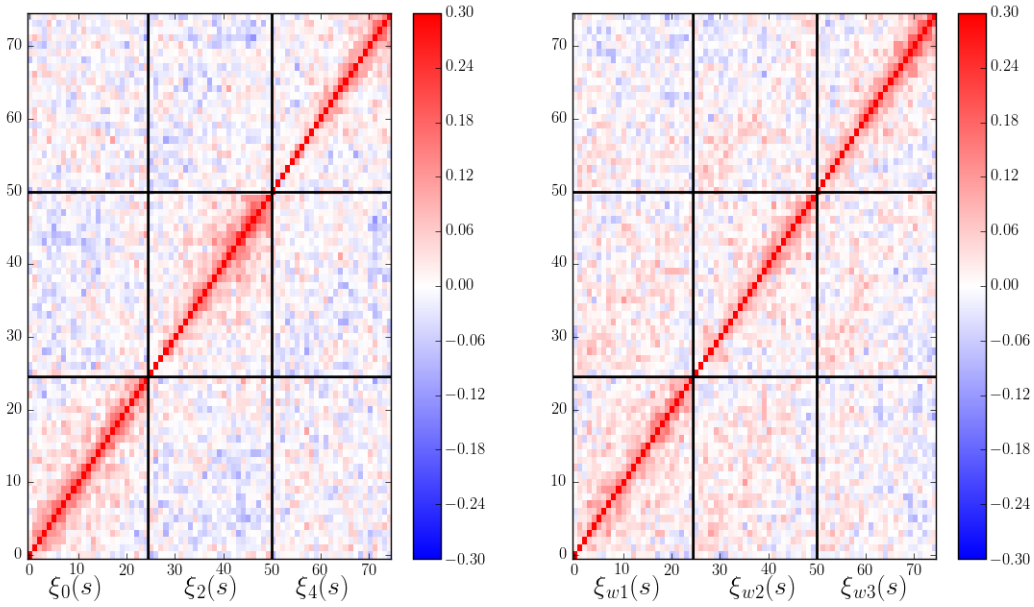


FIGURE 4.4 – Correlation matrices obtained from the 1,000 EZ mocks and used to fit the data for the 3-multipole (left) and 3-wedge (right) analyses. Values of the correlation above 0.3 (along the diagonal) are truncated to enhance the contrast in the lower correlation regions. Each individual square is 25×25 bins of width 8 h^{-1}Mpc from 0 to 200 h^{-1}Mpc.

for a tracer of bias b with a linear growth rate f, assuming a linear power spectrum $P_{\rm lin}$ whose normalization is given by σ_8 :

$$\xi^{\rm CLPT}(\alpha_\parallel s_\parallel, \alpha_\perp s_\perp\,;\, b, f|P_{\rm lin}). \tag{4.9}$$

where $P_{\rm lin}$ is fixed according to the fiducial cosmological parameters we use for the analysis. Here, we have introduced two additional parameters, $\alpha_\parallel$ and $\alpha_\perp$, to account for different dilation of scales for the directions along and perpendicular to the LOS. This approach allows the measured cosmology to differ from the fiducial cosmology from which distances are inferred using redshifts and angular coordinates. The parameters $\alpha_\parallel$ and $\alpha_\perp$ can be related to the expansion rate $H(z)$ and the angular diameter distance $D_{\rm A}$ through :

$$\alpha_\parallel = \frac{H^{\rm fid}(z) r_s^{\rm fid}}{H(z) r_s}, \qquad \alpha_\perp = \frac{D_{\rm A}(z) r_s^{\rm fid}}{D_{\rm A}^{\rm fid}(z) r_s} \tag{4.10}$$

where r_s is the sound horizon at the end of the baryon drag epoch and quantities with the superscript 'fid' refer to quantities determined within the fiducial cosmology. From $\alpha_\parallel$, $\alpha_\perp$, and the fiducial cosmology, one can construct a volume averaged distance, $D_{\rm V}$:

$$D_{\rm V} = \left[(1+z)^2 cz \frac{D_{\rm A}^2}{H} \right]^{\frac{1}{3}} \tag{4.11}$$

where c is the speed of light. One can also define the Alcock-Paczynski parameter $F_{\rm AP}$, which is proportional to the ratio of scales along and perpendicular to the LOS :

$$F_{\rm AP} = \frac{1+z}{c} D_{\rm A} H \tag{4.12}$$

Alternatively, one can also use a combination of $\alpha_\parallel$ and $\alpha_\perp$ such that :

$$\alpha = \alpha_\parallel^{1/3}\alpha_\perp^{2/3}, \qquad \epsilon = (\alpha_\parallel/\alpha_\perp)^{1/3} - 1 \tag{4.13}$$

When using the monopole only, one can constrain the α variable that we often refer to this quantity as $\alpha_{\rm iso}$. It corresponds to an isotropic shift of the BAO feature and gives access to the spherically-averaged BAO distance $D_{\rm V}$ through :

$$\alpha_{\rm iso} = \frac{D_{\rm V}(z) r_s^{\rm fid}}{D_{\rm V}^{\rm fid}(z) r_s} \,. \tag{4.14}$$

For consistency, we also perform an analysis by assuming in the model that there is no anisotropic dilation of scales and fitting $\alpha_{\rm iso}$.

[13] derived an analytical formula for this parameter defined by $\alpha_{\rm iso}^{m+n} = \alpha_\parallel^m \alpha_\perp^n$, where m and n provide the degree of degeneracy in the parallel and perpendicular to the LOS directions. He computed predictions for different configurations : when using the monopole only (green dashed line), when removing the effect of RSD on the monopole (black solid line) and using the quadrupole too (red dotted line). Figure 4.5 shows the EZ mocks measurements on $\alpha_\parallel$ and $\alpha_\perp$ in Fourier space [14] compared to the analytical predictions from [13]. The mock results display a degenerate direction on $\alpha_{\rm iso}^{m+n} = \alpha_\parallel^m \alpha_\perp^n$ which is close to the one predicted when using the monopole only with the RSD removal which corresponds to $m = 1/3$ and $n = 2/3$. So, given the statistical precision of the current sample, we can consider that $\alpha_{\rm iso} \simeq \alpha_\parallel^{1/3}\alpha_\perp^{2/3}$ is a valid approximation.

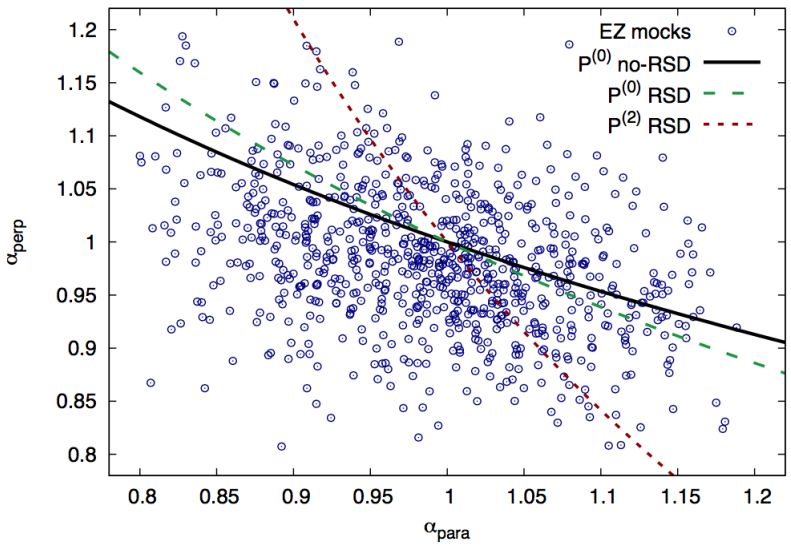


FIGURE 4.5 – EZ mock measurements for $\alpha_\parallel$ and $\alpha_\perp$ (blue symbols) along with the degeneracy directions predicted by Ross et al. 2015 when only the monopole is used after (black solid line) and before (green dashed line) the RSD are removed, and when the quadrupole is used (red dotted line). From [14].

4.1.4 Parameter inference

We extract the results of the fitting of either the three first Legendre multipoles or the three wedges by minimizing the χ^2 defined by :

$$\chi^2 = (\xi^{\text{Data}} - \xi^{\text{Model}})C^{-1}(\xi^{\text{Data}} - \xi^{\text{Model}})^T \tag{4.15}$$

where ξ^{Data} corresponds to the measurement, ξ^{Model} to the associated theoretical prediction, and C^{-1} the inverse of the estimated covariance matrix. The latter includes corrections due to finite number of mocks and number of bins in the analysis that can bias the measurements. We follow the procedure described in [15] and referred as the Hartlap correction :

$$C^{-1}_{\text{unbiased}} = (1 - D)C^{-1}_{\text{mock}} \text{ with } D = \frac{N_b + 1}{N_m - 1} \tag{4.16}$$

where N_b is the total number of bins in the measurements (in this analysis, the reference uses 15 bins) and N_m is the number of realizations. Thus, the smaller the number of realizations N_m, the larger these corrections are. In addition, the effect of noise propagates to the parameters constraints, so that the obtained variance of each parameter needs to be rescaled by [16] :

$$M = \sqrt{\frac{1 + B(N_b - N_p)}{1 + A + B(N_p + 1)}} \tag{4.17}$$

where N_p is the number of fitting parameters and the two factors A and B are :

$$A = \frac{2}{(N_m - N_b - 1)(N_m - N_b - 4)} \tag{4.18}$$

$$B = \frac{(N_m - N_b - 2)}{(N_m - N_b - 1)(N_m - N_b - 4)} \tag{4.19}$$

The following table presents the correction factors for each set of synthetic mock catalogs, with $N_b = 3 \times 15$ bins for the measurements from the three multipoles or wedges and $N_p = 5$ fitting parameters :

mock catalogs	N_m	$1 - D$	M
QPM mocks	400	0.96	1.005
EZ mocks	1000	0.98	1.004

The Hartlap correction represent $2 - 4\%$ factor in the χ^2 values and the ones described in [16] have a minor contribution to the final errors.
In the fitting, the covariance matrices are determined from the EZ mocks with a correction to equalize small differences in area. Eventually, we find the χ^2 minima using the MINUIT libraries [2]. Error-bars are derived from the $\Delta\chi^2 = 1$ region of the marginalized χ^2 profiles and are allowed to be asymmetric.

In section 5.2, we will will compare our results with the companion papers, we also run Markov-chains to compute the likelihood surface of the set of parameters. We use the *emcee* package [17] which is a python implementation of the affine-invariant ensemble sampler for Markov chain Monte Carlo (MCMC) ; we check its convergence using the Gelman-Rubin convergence test with the condition $R - 1 < 10^{-2}$.

2. James, F. MINUIT Function Minimization and Error Analysis : Reference Manual Version 94.1. 1994.

4.2 Modelling systematics : study on Outer Rim

Given that all perturbative approaches are approximate methods to solve the dynamics of gravitational clustering, it is necessary to test the domain of validity of the theoretical predictions using numerical simulations. In practise, this is how the fitting range of each RSD model is fixed. [18] (resp. [19, 20]) tested the GS (resp. CLPT-GS) model using a set of N-body simulations presented in [21] for halos of the appropriate mass range to host BOSS galaxies at redshift $z \simeq 0.5$. The CLPT-GS model has also been applied on BOSS DR11 data by [22] and on BOSS DR12 by [23]. In particular, [20] showed that the difference between the CLPT-GS model and N-body results for the matter density field was less than 5% for the Legendre multipoles in redshift space for scales above $\simeq 20h^{-1}$Mpc.

4.2.1 Production of the Outer Rim mock catalogs

In this work, we resort to the N-body Outer Rim simulation presented in section 3.2.1. We briefly remind the most important properties. The volume of the Outer Rim simulation is a cube of side $L = 3000\,h^{-1}$Mpc with 10240^3 dark matter particles. The mass resolution of each dark matter particle is $m_p = 1.82 \times 10^9\,h^{-1}\mathrm{M}_\odot$. The mass resolution is enough to resolve halos that host quasars in the eBOSS redshift range. However, the volume is not enough to cover the full footprint of the DR14 quasar sample, even with the remapping technique. The cosmological parameters are given in table 4.1 and are consistent with the WMAP7 cosmology [24]. The initial conditions are calculated at z=200 using the Zel'dovich approximation.

We build the mocks from a single snapshot at $z = 1.433$ for which halos of more than 20 particles are available. In a first approach (dubbed "mass bin"), we consider that only halos with mass $M = 10^{12.5\pm0.3}\,h^{-1}\mathrm{M}_\odot$ can host a quasar. In a refined approach, we apply the (5+1)-parameter HOD using the parameters derived for the QPM mocks and defined by :

For each halo, we determine the concentration from the halo mass using an ad-hoc parameterization of the data described in [25]. The position of the satellites and their velocity are drawn from a profile according to the NFW prescription [26]. We can vary the fraction of satellites by playing with the mass of the biggest halos that contain substructures. For instance, the fraction of satellites can be increased to $f_{\mathrm{sat}} = 25\%$ by setting $\mathrm{M}_{\mathrm{sat}} = 10^{15}\,h^{-1}\mathrm{M}_\odot$ in the HOD model. In addition, we take advantage of the fact that the eBOSS quasar measurement is shot-noise dominated, and that the duty cycle for quasars is low ($\tau = 1.2\%$), to draw many realizations (up to 100) from the same parent box.

The production of the Outer Rim mock catalogs for the DR14 quasar analysis required some modifications compared to the existing MAKE SURVEY implementation. We briefly describe these modifications step by step :

(i) **Volume remapping** : we found some issues related to the remapping technique for the Outer Rim simulation that were not observed with the QPM boxes. Figure 4.6 shows the monopole of the correlation function in real space for Outer Rim (left) and QPM (right) simulations and for different configurations : when just applying the sky projection to translate cartesian coordinates (x,y, z) into (RA,DEC,redshift) using the original code (yellow) and our PYTHON version (cyan), when applying the remapping step before (green and pink). The dashed line shows the theoretical prediction from CLPT. We can see that all the configurations yields the same clustering

for the QPM, while differences appear at large scales when applying or not the remapping technique for the Outer Rim. The issue could be related to the box size of the Outer Rim simulation (3 Gpc) compared to the one of QPM (5 Gpc). A smaller volume implies that some large k-modes will not be probed by the simulation which can affect the clustering when we move parts of the box to another location. We do not apply the remapping technique and consider cubic boxes only.

(ii) **Translation / rotation** : we did not apply any translation or rotation, we worked with the original cubic box.

(iii) **Sky projection** : the cartesian coordinates are transformed in right ascension, declination and redshift using Outer Rim cosmology (see table 4.1).

(iv) **Apply redshift space distortions** : The implementation of RSD is made on the redshift where $z_{\rm red} = z_{\rm real} + dz$ with

$$dz = \frac{v_z}{c}(1 + z_{\rm real}) \tag{4.20}$$

where the z-axis is taken to be the LOS axis. However, we have just one snapshot at $z_{\rm sim} = 1.433$ to cover the redshift range $0.8 \leq z \leq 2.2$. Both methods are approximate and additional snapshots are required to build a light-cone mock that includes the redshift evolution of the cosmological fields (for instance, the EZ mocks were built from 7 snapshots).

(v) **Apply the survey footprint** : the cubic box of 3 Gpc side is not big enough to include the eBOSS DR14 quasar survey footprint. Angular cuts are applied instead and an area of 1888 deg^2 with uniform redshift coverage can be selected.

(vi) **Randomly downsample based on sky completeness** : we did not apply this step for the eBOSS DR14 quasar analysis because it would require applying the survey footprint first.

(vii) **Randomly downsample based on the radial selection** to ensure that the mock catalogs have the same redshift distribution as the data (applied). However, since the Outer Rim mocks catalogs have been created from a single snapshot at $z = 1.433$, and since we are just interested in evaluating the performance of the RSD model, we apply an additional redshift cut $z \leq 2.0$ to produce an effective redshift that matches the one of the single snapshot. It also allows us to compare, at the same effective redshift, the real space results using directly the output of the Outer Rim simulation and the results in redshift space after applying the procedure we have just described.

4.2.2 Validity of CLPT in real space

The linear power spectrum corresponding to the Outer Rim cosmology given in table 4.1 is given as input to the CLPT code which calculates the real space observables $\xi(r)$, v_{12}, σ_{12} as function of the first two Lagrangian parameters. The dependency on the bias parameters is given by equation 6.12. We then compare the theoretical predictions for these obserables in real space with results from the Outer Rim simulation.

Figure 4.7 presents the agreement in real space between the Outer Rim results (blue points) for the catalog without satellites and the CLPT prediction (red) compared to the linear theory prediction (green) for the real-space correlation function, the mean infall pairwise velocity, and the velocity dispersion. The magenta curve in the top panel corresponds to the quadrupole of the correlation function in real space which is compatible

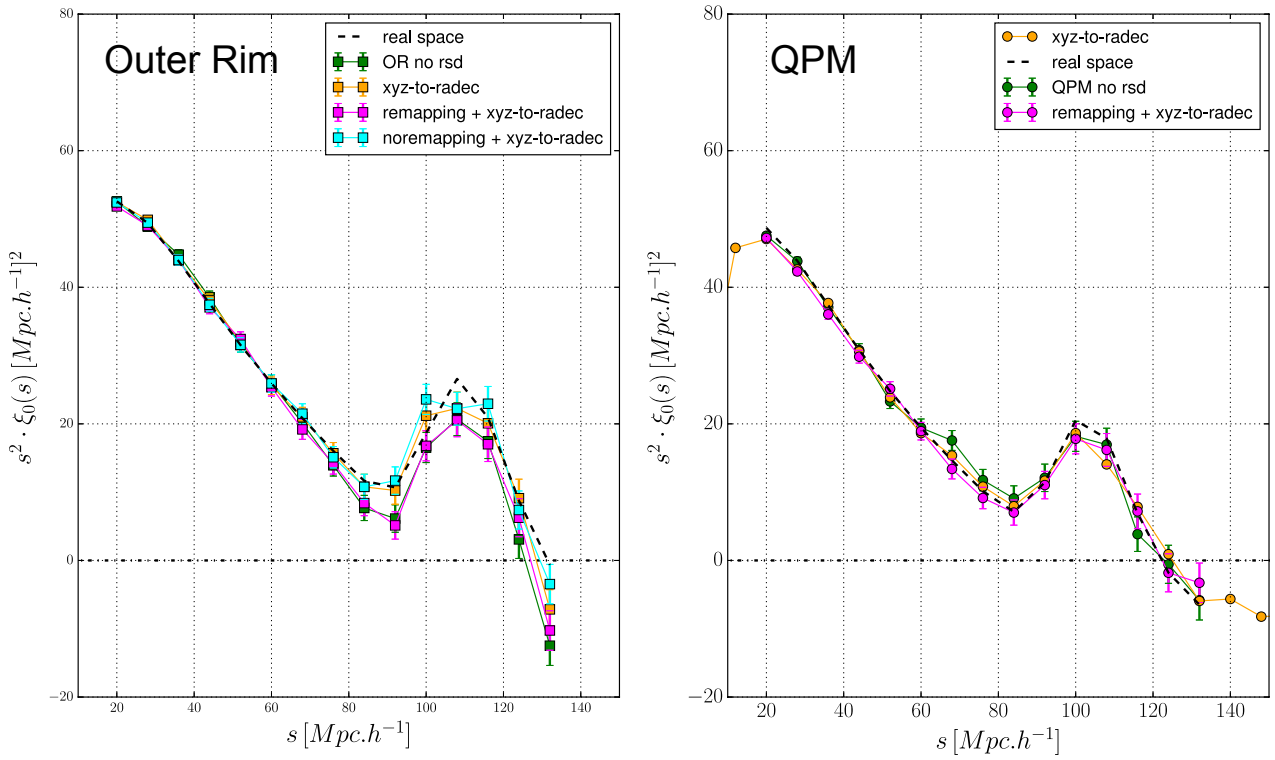


FIGURE 4.6 – Monopole of the correlation function in real space for the Outer Rim (left) and QPM (right) simulations and for different configurations with or without applying the remapping technique.

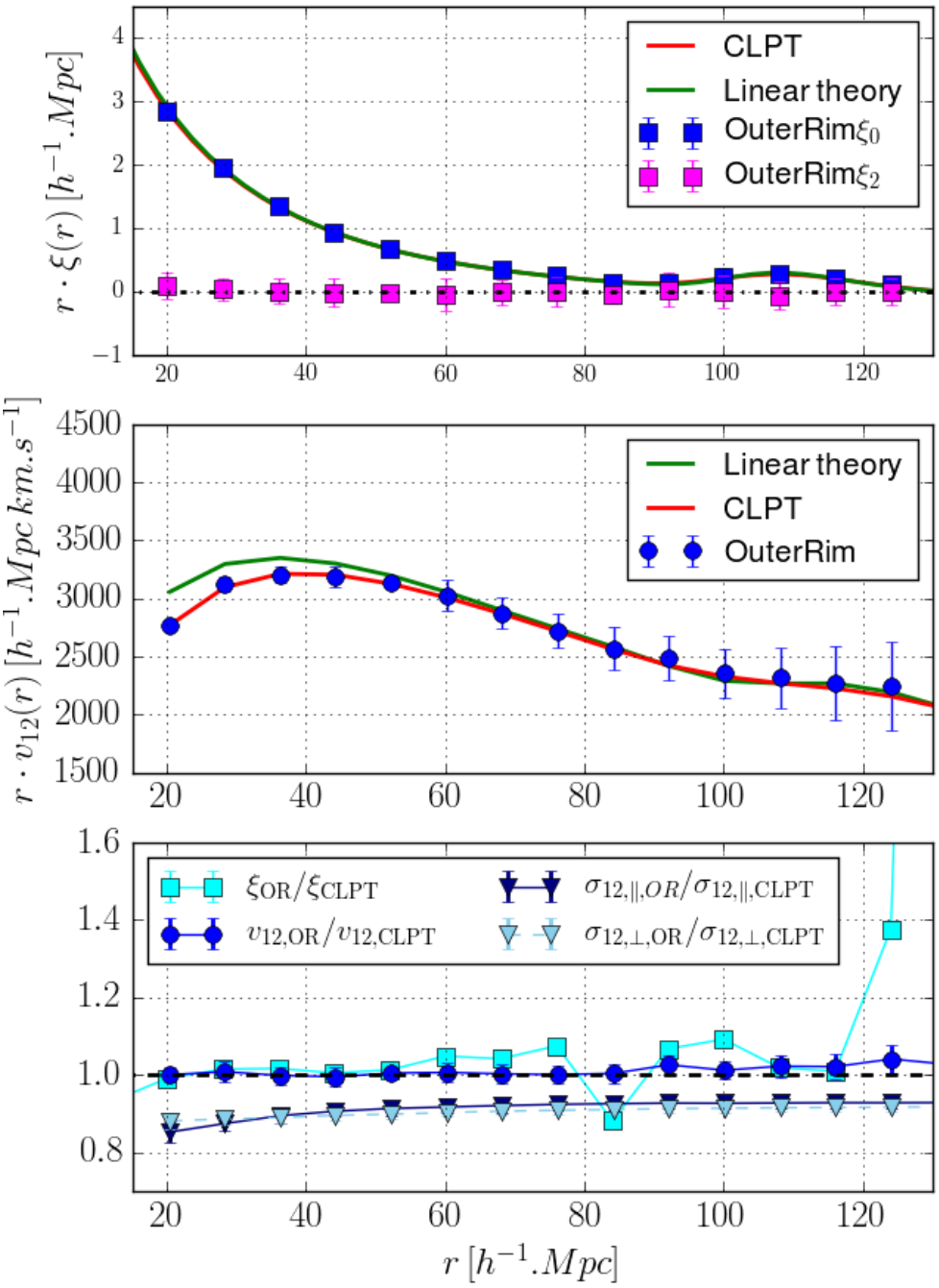


Figure 4.7 – Real space observables for the case without satellite and redshift smearing. Blue points correspond to the results from the Outer Rim simulation, the red curve is the CLPT prediction and the green one is the linear theory prediction. Top panel : correlation function in real space. Middle panel : pairwise infall velocity. Bottom panel : Ratio between Outer Rim results and CLPT predictions for ξ (square), v_{12} (circle) and $\sigma_{12,\parallel,\perp}$ (triangles).

with zero in the N-body simulation as expected in real space where the redshift coordinate accurately measures the distance along the LOS. However, the radial distances that are inferred from redshifts obtained in spectroscopic surveys also include components from peculiar velocities which give rise to the anisotropic clustering we observe, and thus to a non-zero quadrupole of the correlation function in redshift space. The bottom panel displays the ratio between the results from the Outer Rim simulation and the CLPT predictions. We confirm that at the mean redshift of eBOSS quasar sample, $z \simeq 1.5$, CLPT reproduces well the clustering and velocity statistics in real space for halos of masses of the order of $10^{12.5}\mathrm{M}_{\odot}$ on scales of interest (above $\simeq$20 h^{-1}Mpc) which is in agreement with the [20] determination for ranges of halo masses that correspond to the BOSS LRG clustering at $z \simeq 0.5$.

Using the assumption of a linear coupling between the matter density field and the tracer velocity field , we remind that one can derive the linear theory prediction for the pairwise mean infall velocity :

$$v_{12}(r) = -r\frac{fb}{\pi^2}\int kP_m^r(k)j_1(kr)dk \tag{4.21}$$

where $j_1(kr)$ is the first-order spherical Bessel function. We can see on the middle panel of figure 4.7 that the pairwise mean infall velocity measured on the N-body simulation deviates from the linear theory (green curve) on scales below $\sim$60 h^{-1}Mpc. Given that it is directly proportional to the growth of cosmic structure, providing reliable cosmological constraint on this parameter requires precise modeling of the non-linear evolution of the matter and density fields. A similar prediction can be derived for σ_{12} and which is given by equation 3.112 (both derivations can be found in [18]). In this expression, the velocity field is assumed to be unbiased w.r.t. the matter density field. The effect of velocity bias was studied in [27] and is expected to be of the order of a few percent for f.

Equation 4.21 shows that the mean infall velocity is expected to be proportional to the bias and to the linear growth rate of structure on large scales. Since $\xi_{\mathrm{tracer}} = b^2\xi_{\mathrm{m}}$, another interesting test in real space is to check that the same bias value can reproduce both correlation function and infall velocity. Figure 4.7 presents results for $F' = 1.33$ which corresponds to $b\sigma_8 = 0.990$; this value is consistent with bias measurements in redshift space for the case without satellite and redshift smearing. The bottom panel of figure 4.7 also shows that the velocity dispersion terms parallel (dark blue triangles) and perpendicular (light blue triangles) to the separation of the pair present a 10% offset compared to CLPT predictions. This issue has previously been discussed in [18] and [20]. We will see in section 4.2.4 that adding a constant shift to the CLPT predictions to match the velocity dispersion observed in Outer Rim does not affect the cosmological parameters when fitting on observables in redshift space.

4.2.3 Performance of the CLPT-GS model in redshift space

In this section, we investigate the response of the RSD model by fitting the redshift space correlation function of the mocks created from the Outer Rim simulation and by comparing the cosmological parameters to the expected values ($f\sigma_8 = 0.382$ and $\alpha_{||} = \alpha_{\perp} = 1$). When not specified, the reference uses a covariance matrix from the NGC EZ mocks without adding close-pairs or redshift failures treatment and is rescaled to match the statistics of the Outer Rim catalogs. F'' is fixed according to the peak-background split assumption using the ST mass function, and the fit uses data from 16 h^{-1}Mpc to 138 h^{-1}Mpc with bin width of 8 h^{-1}Mpc. The results of the fits are presented in Table 4.2.

Different sets of Outer Rim catalogs have been produced in order to study the effect of satellite fraction and redshift uncertainties :

— **Satellite fraction** : the presence of quasars hosted in satellite halos increases the amount of virialized objects within a halo ; this increase modifies the small-scale clustering as it corresponds to a strong elongation of structures along the LOS known as the Fingers-of-God (FoG) [28] effect. It also affects the amplitude of the clustering at all scales because of the dependence of the number of satellites with the mass of the halos. We study the case with $f_{\rm sat} = 13\%$ satellite fraction as implemented in the QPM mocks but also the cases $f_{\rm sat} = 0\%$ and $f_{\rm sat} = 25\%$ for systematics checks.

— We apply two different distributions for **redshift uncertainties** : a Gaussian redshift distribution according to the SRD and a physical redshift distribution according to the distribution $(z_{\rm MgII} - z)$ as seen in the data in Figure 2.17. For the latter, we rescale the distribution so that the width matches the one of the SRD in order to focus the study on the effect of the exponential tails in the observed distributions.

In Table 4.2, we report the results for the 3-multipole and 3-wedge analyses where small systematic shifts between the two methods can be observed at the level of $\Delta f\sigma_8 = 0.006$, $\Delta\alpha_{\parallel} = 0.005$ and $\Delta\alpha_{\perp} = 0.006$.

4.2.3.1 Quasar-halo connection

Figure 4.8 compares the multipoles (top panel) and the wedges (bottom panel) of the correlation function for the “mass bin" (only halos with mass $M = 10^{12.5\pm0.3} M_{\odot}$ can host a quasar) and the HOD biasing scenarios. For the latter, the data points are obtained from the average of 100 realisations. At the largest scales shown, the results from Outer Rim tend to deviate from the predictions of the model in a region where these predictions do not differ from linear theory. It may be due to the simulation box size effects, but we do not use scales larger than 138 h^{-1}Mpc in our fit range and the deviation is much smaller than the statistical precision of the data. For the monopole, it is clear that the “mass bin" scenario is better reproduced by the model at all scales and that, in the region of the BAO feature, the HOD presents an unexpected behaviour. Therefore, with the present version of these mocks we may anticipate differences in the extracted geometrical parameters $\alpha_{\parallel}$ and $\alpha_{\perp}$. Furthermore, Figure 4.8 reveals the impact of a $\pm10\%$ variation of the parameter $F_{\rm AP}$ (green band) and a $\pm10\%$ variation of $f\sigma_8$ (grey band), showing that the quadrupole is equally sensitive to variations of $F_{\rm AP}(\propto \alpha_{\perp}/\alpha_{\parallel})$ and $f\sigma_8$. But it also demonstrates that the hexadecapole is mostly sensitive to the variations of the geometrical parameters and hence will contribute to break this degeneracy. As expected, for the wedges, since the sum of the three wedges corresponds to the monopole, the effect is more degenerate among the three wedges and the wedge in the middle is the least affected as it probes pairs with intermediate angles between parallel and perpendicular to the LOS.

For the HOD case, we varied the satellite fraction and present the measured monopole (blue), quadrupole (red) and hexadecapole (green) obtained in the top panel of figure 4.9. Increasing the satellite fraction mildly enhances the amplitude of the clustering, and the quadrupole and hexadecapole are almost unaffected. While no large difference between satellite fractions is seen in the mocks, previous analyses of the data tend to favour a satellite fraction around 0.15. This behaviour is shown in figure 9 of [29] which compares the projected quasar correlation function measurements to the HOD model we use in QPM mocks and in the OuterRim for the case $f_{sat} = 0.15$. The exact satellite fraction for the halos hosting quasars however, is not known precisely, and is degenerate with the

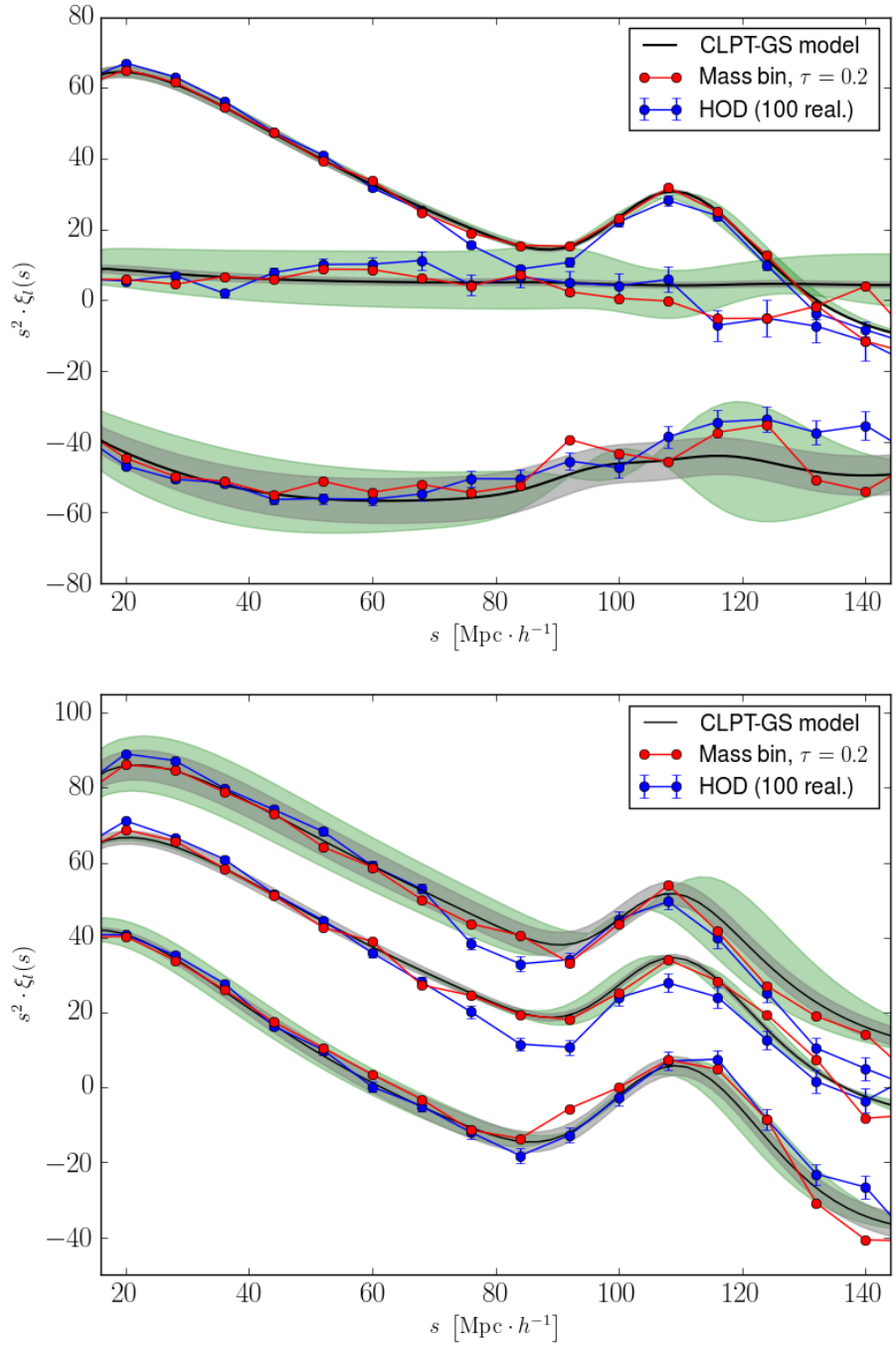


FIGURE 4.8 – Top panel : Monopole of the correlation functions for the two bias models considered : “mass bin" (red) and HOD (blue). For the HOD the data points are obtained from the average of 100 realisations. The CLPT model has been adjusted on the “mass bin" points (solid line). The green band shows the effect of a ±10% variation of the parameter $F_{\rm AP}(\propto \alpha_\perp/\alpha_\parallel)$ and the grey band shows the effect of a ±10% variation of $f\sigma_8$. Bottom panel : Same for the three wedges.

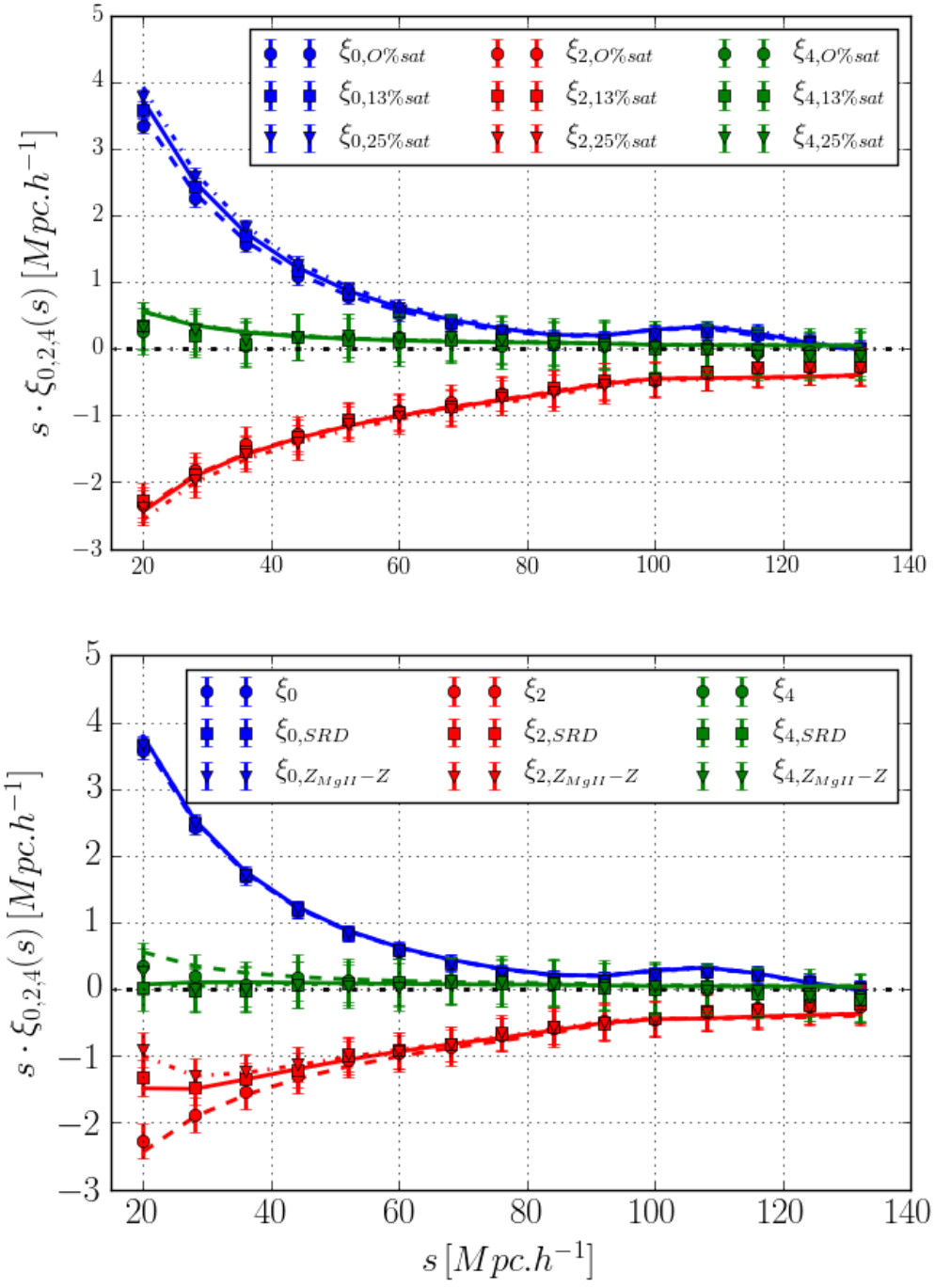


FIGURE 4.9 – Top panel : Monopole (blue), quadrupole (red) and hexadecapole (green) for 3 satellite fractions without redshift smearing with the model set to the best fitting parameters for 0% satellite (dashed line), 13% satellite (solid line) and 25% satellite (dashdot line). Bottom panel : Monopole (blue), quadrupole (red) and hexadecapole (green) for 3 redshift smearing and 13% satellite with the model set to the best fitting parameters for no smearing (dashed line), SRD smearing (solid line) and ($z_{\rm MgII} - z$) smearing (dashdot line).

duty cycle of quasars that probably varies with luminosity and redshift. We therefore report the average value obtained for the three satellite fractions (0%,13% and 25%) in table 4.2 when estimating the systematic error related to the modeling. Further studies to constrain the dark matter halo mass and duty cycle using the final eBOSS quasar clustering measurements would provide a superior statiscal power to investigate these effects, following the approach developed in [30] and [11].

4.2.3.2 Redshift uncertainties

We first investigate the response of the model when no smearing due to redshift error is applied. For all the cases considered, we observe a systematic shift of $f\sigma_8$ towards lower values and the maximum offset w.r.t the input cosmology is $\Delta f\sigma_8 = -0.014$. For $\alpha_\parallel$, the maximum offsets for the HOD ($\Delta\alpha_\parallel = 0.038$) is much larger than for the 'mass bin' ($\Delta\alpha_\parallel = 0.016$). This situation probably arises from the difference observed on the monopole : the HOD brings more high-mass halos whose velocity dispersion in N-body simulations is actually lower than that of low-mass halos. This is related to the environment probed by halos, where high-mass halos correspond to more specific environments. The fact we observe differences between two ways of populating the halos demonstrates the need for a better understanding of the impact of the astrophysics conditions leading to the formation of quasars. For $\alpha_\perp$, the results are consistent with an offset smaller than $\Delta\alpha_\perp = 0.006$. All these estimates receive contributions from both the biasing scenarios and from the modeling of the correlation function, but presently they should be viewed as global intrinsic systematic errors in our measurement.

The impact of redshift resolution is studied either by drawing the redshift from a Gaussian distribution according to eBOSS SRD (solid lines) or by drawing the redshift from the "physical" distribution of $(z_{\rm MgII} - z)$ as shown in figure 2.17. The bottom panel of figure 4.9 reveals that, for the two types of smearing, the quadrupole and the hexadecapole are affected at scales below $\sim$50 h^{-1}Mpc and that the monopole is unaffected. It also shows that applying a more physical smearing with exponential tails has a larger effect on the quadrupole. To account for redshift smearing in the RSD modeling, we add a constant dispersion velocity term to the width of the Gaussian distribution used for $\mathcal{G}$ in equation 6.10 following the approach in [31] :

$$\sigma_{12}^2(r,\mu) = \sigma_{12,\rm CLPT}^2(r,\mu) + \sigma_{\rm tot}^2 \,. \tag{4.22}$$

This additional term can be decomposed as $\sigma_{\rm tot}^2 = \sigma_{\rm FoG}^2 + \sigma_z^2$ where $\sigma_{\rm FoG}$ is produced by the Finger-of-God effect due to virialized motions of the quasars within their host halo and σ_z arises from the smearing due to redshift resolution. However, the two parameters are degenerate, and in the model a single total nuisance parameter is used to represent this effect.

A sizeable effect of the redshift smearing is observed for the parameter $f\sigma_8$ extracted from the fits as presented in table 4.2. For the cases considered, an average systematic shift of $\Delta f\sigma_8 \simeq -0.010$ exists for the SRD smearing and an effect of $\Delta f\sigma_8 \simeq -0.021$ for the physical redshift smearing. This systematic shift could, in principle, be reduced by using the actual shape of the redshift error distribution in a future modified streaming model. For $\alpha_\parallel$, there is a small compensation of the large effect seen for the HOD when applying the SRD smearing which is slightly reduced when using the physical smearing. For the "mass bin", a similar behaviour is observed but remains smaller that for the HOD. No effect is seen on $\alpha_\perp$.

In summary, the overall systematic shifts due to the modeling of the full-shape anisotropic correlation function with our CLPT-GS model are $\Delta f\sigma_8 = 0.033$, $\Delta\alpha_{\parallel} = 0.038$, and $\Delta\alpha_{\perp} = 0.006$ where we use the maximum deviation observed for the two biasing scenarios and the two redshift smearing options.

Table 4.3 presents the results for more restrictive hypotheses on the cosmology where the cosmology is either fixed to the input cosmology of OuterRim or where we allow for an isotropic variation of the geometrical parameters, namely $\alpha_{\parallel} = \alpha_{\perp} = \alpha_{\rm iso}$. This test is performed using the physical redshift smearing in the case of the multipole analysis ; for the HOD, we list the results for the set with $f_{\rm satt} = 13\%$ that is favoured by the data. In these conditions, the systematic shift on $f\sigma_8$ w.r.t. the input cosmology is reduced to $\Delta f\sigma_8 = 0.017$; for the parameter $\alpha_{\rm iso}$, the maximum variation observed among all mocks that were produced is $\Delta\alpha_{\rm iso} = 0.024$.

TABLE 4.2 – Impact on measured cosmological parameters for the different halo populating approaches and redshift smearing options. For the input cosmology $f\sigma_8 = 0.382$ and $\alpha_{\parallel} = \alpha_{\perp} = 1$.

config	smearing	$b\sigma_8$	$f\sigma_8$(OR = 0.382)	$\alpha_{\parallel}$(OR = 1.0)	$\alpha_{\perp}$(OR = 1.0)	$\sigma_{\rm tot}$
3-multipole						
HOD	no	1.024 ± 0.001	0.377 ± 0.002	1.031 ± 0.002	1.001 ± 0.001	1.026 ± 0.1
HOD	SRD	1.024 ± 0.001	0.363 ± 0.002	1.021 ± 0.002	1.005 ± 0.001	5.48 ± 0.03
HOD	$(z_{\rm MgII} - z)$	1.028 ± 0.002	0.355 ± 0.003	1.028 ± 0.003	0.998 ± 0.001	6.73 ± 0.03
mass bin	no	0.966 ± 0.005	0.377 ± 0.006	1.014 ± 0.006	1.002 ± 0.005	1.26 ± 0.130
mass bin	SRD	0.971 ± 0.005	0.368 ± 0.006	1.011 ± 0.007	1.002 ± 0.005	5.60 ± 0.040
mass bin	$(z_{\rm MgII} - z)$	0.976 ± 0.005	0.355 ± 0.007	1.025 ± 0.008	0.994 ± 0.005	6.84 ± 0.036
3-wedge						
HOD	no	1.025 ± 0.001	0.368 ± 0.003	1.038 ± 0.002	0.995 ± 0.002	1.58 ± 0.1
HOD	SRD	1.029 ± 0.001	0.360 ± 0.003	1.025 ± 0.003	1.003 ± 0.002	5.42 ± 0.03
HOD	$(z_{\rm MgII} - z)$	1.031 ± 0.001	0.353 ± 0.003	1.025 ± 0.003	1.002 ± 0.002	6.55 ± 0.03
mass bin	no	0.968 ± 0.005	0.372 ± 0.007	1.016 ± 0.007	1.001 ± 0.005	1.34 ± 0.130
mass bin	SRD	0.974 ± 0.005	0.362 ± 0.008	1.012 ± 0.008	1.002 ± 0.006	5.58 ± 0.043
mass bin	$(z_{\rm MgII} - z)$	0.979 ± 0.005	0.349 ± 0.008	1.022 ± 0.008	0.996 ± 0.006	6.57 ± 0.042

4.2.4 Additional tests

Finally, we perform a series of tests for the mocks on the “mass bin" case with SRD redshift smearing to study the impact of the ingredients of the model on the cosmological parameters. In particular, we examine the following effects whose results are summarized in table 4.4 :

— F'' : in the reference, F'' is fixed under the peak-background split assumption using the ST mass function and the result of the fit using the PS mass function is the same. When F'' is set free in the fit, small changes in the cosmological fit parameters are observed and are compatible with the variations of the statistical errors. The nuisance parameter $\sigma_{\rm tot}$ and its error are affected, suggesting a probable degeneracy with F''. There is also an effect on the linear bias parameter with a shift $\Delta b\sigma_8 = 0.037$. We therefore do not report any bias measurement in the final cosmological results for this sample ; further investigations on the bias models and prescriptions are needed for the final sample if we want to constrain the astrophysical properties of quasars using bias measurement from full-shape analysis.

— $\sigma_{\rm tot}$: When fixing $\sigma_{\rm tot} = 5.7h^{-1}$Mpc (i.e., the average value of the SRD resolution used to create the mocks), the cosmological parameters of the simulation are

recovered and the precision on $\alpha_\parallel$ is improved by 30%. This result is achieved because, when fixing $\sigma_{\rm tot} = 5.7h^{-1}$Mpc, the small scale statistical power is available for constraining $\alpha_\parallel$. Although this result should be viewed as a consistency check only, it demonstrates that a better knowledge of the redshift precision is important for the analysis of the full eBOSS quasar sample.

— $r_{\rm min}$: Setting the lower bound of the fit range to $r_{\rm min} = 24\ h^{-1}$Mpc instead of 16 h^{-1}Mpc produces almost no variation of $f\sigma_8$ and an effect on $\alpha_\parallel$ which is within the statistical precision.

— $\sigma^2_{\rm shift}$: Adding a small constant shift to the CLPT predictions for the velocity dispersion to match the one observed in the OuterRim simulation in real space produces no significant effect on the measured cosmological parameters. This constant shift is much smaller than the redshift resolution uncertainties : $\sigma^2_{\rm shift} = 3\,h^{-1}$.Mpc while $\sigma^2_z \sim 6^2\,h^{-1}$.Mpc.

TABLE 4.3 – Comparison between different hypotheses on the cosmology : cosmology is fixed to the input of Outer Rim ($\alpha_\parallel = \alpha_\perp = 1.$), isotropic case ($\alpha_\parallel = \alpha_\perp = \alpha_{\rm iso}$), and anisotropic case ($\alpha_\parallel$ and $\alpha_\perp$). Results are given for physical redshift smearing and for the 3-multipole analysis

config	cosmology	$b\sigma_8$	$f\sigma_8$	$\alpha_\parallel$	$\alpha_\perp$	$\sigma_{\rm tot}$
mass bin	Outer Rim	0.961 ± 0.005	0.370 ± 0.005	fixed	fixed	$6.40^{+0.26}_{-0.27}$
mass bin	isotropic	0.966 ± 0.005	0.371 ± 0.006	$\alpha_{\rm iso} = 1.005 \pm 0.005$	–	$6.44^{+0.28}_{-0.29}$
mass bin	anisotropic	0.976 ± 0.005	0.355 ± 0.007	1.025 ± 0.008	0.994 ± 0.005	$6.84^{+0.36}_{-0.36}$
HOD $f_{\rm sat} = 13\%$	Outer Rim	1.015 ± 0.001	0.368 ± 0.002	fixed	fixed	$6.50^{+0.82}_{-1.07}$
HOD $f_{\rm sat} = 13\%$	isotropic	1.022 ± 0.002	0.366 ± 0.002	$\alpha_{\rm iso} = 1.005 \pm 0.001$	-	$6.37^{+0.89}_{-1.07}$
HOD $f_{\rm sat} = 13\%$	anisotropic	1.027 ± 0.002	0.351 ± 0.003	1.027 ± 0.002	0.994 ± 0.002	$6.73^{+1.38}_{-1.44}$
HOD $f_{\rm sat} = 0\%$	Outer Rim	0.980 ± 0.001	0.363 ± 0.002	fixed	fixed	$6.05^{+1.07}_{-1.40}$
HOD $f_{\rm sat} = 0\%$	isotropic	0.999 ± 0.002	0.367 ± 0.002	$\alpha_{\rm iso} = 1.017 \pm 0.001$	-	$6.08^{+1.10}_{-1.38}$
HOD $f_{\rm sat} = 0\%$	anisotropic	1.006 ± 0.002	0.355 ± 0.003	1.035 ± 0.003	1.010 ± 0.002	$6.35^{+1.63}_{-1.73}$
HOD $f_{\rm sat} = 25\%$	Outer Rim	1.046 ± 0.001	0.376 ± 0.002	fixed	fixed	$6.48^{+0.72}_{-0.90}$
HOD $f_{\rm sat} = 25\%$	isotropic	1.050 ± 0.002	0.374 ± 0.002	$\alpha_{\rm iso} = 1.002 \pm 0.001$	-	$6.39^{+0.77}_{-0.88}$
HOD $f_{\rm sat} = 25\%$	anisotropic	1.052 ± 0.002	0.360 ± 0.002	1.020 ± 0.002	0.990 ± 0.002	$6.72^{+1.04}_{-1.14}$

TABLE 4.4 – Additional tests performed when varying hypotheses on the second order bias parameter F'', on the total velocity dispersion $\sigma_{\rm tot}$, and on the lower bound of the fit range $r_{\rm min}$. These tests are performed for the "mass bin" case with Gaussian redshift smearing and for the multipole analysis.

config	hypothesis	$b\sigma_8$	$f\sigma_8$	$\alpha_\parallel$	$\alpha_\perp$	$\sigma_{\rm tot}$
mass bin SRD	ref : uses $F'' = F''_{\rm ST}$	0.971 ± 0.005	0.368 ± 0.006	1.011 ± 0.007	1.002 ± 0.005	$5.60^{+0.38}_{-0.40}$
mass bin SRD	$F'' = F''_{\rm PS}$	0.969 ± 0.005	0.369 ± 0.007	1.011 ± 0.007	1.003 ± 0.005	$5.49^{+0.40}_{-0.42}$
mass bin SRD	$F''_{\rm free} = -3.461^{+1.803}_{-1.239}$	0.934 ± 0.009	0.376 ± 0.007	1.001 ± 0.007	1.000 ± 0.005	$4.04^{+1.05}_{-1.12}$
mass bin SRD	$\sigma_{\rm tot} = 5.7\ h^{-1}$Mpc	0.969 ± 0.005	0.369 ± 0.007	1.019 ± 0.005	1.000 ± 0.005	fixed
mass bin SRD	$\sigma_{\rm shift}$	0.967 ± 0.005	0.370 ± 0.006	1.010 ± 0.006	1.003 ± 0.005	$5.71^{+0.38}_{-0.40}$
mass bin SRD	$r_{\rm min} = 24\ h^{-1}$Mpc	0.948 ± 0.006	0.369 ± 0.007	0.998 ± 0.008	0.997 ± 0.005	$5.18^{+0.93}_{-1.09}$

In light of this study, we conclude that the CLPT-GS model can be used for the clustering analysis of the eBOSS quasar sample at $0.8 \leq z \leq 2.2$ with overall systematic errors due to the modeling of the full-shape anisotropic correlation function of :

$$\Delta f\sigma_8 = 0.033 \quad \Delta\alpha_\parallel = 0.038 \quad \Delta\alpha_\perp = 0.006 \tag{4.23}$$

The systematic errors between the 3-multipole and 3-wedge methods are similar and the errors reported are always the largest of the two possibilities. For the analysis of the final eBOSS quasar sample, further work on improving the fidelity of the Outer Rim-based mocks and understanding the difference in the bias models is needed. In particular, [32] extended the CLPT-GS formalism to take into account contributions from Effective Field Theory (EFT) and additional bias terms. They showed that the effects of the biasing scheme are as important as higher-order corrections to the theoretical predictions. Therefore it would be interesting to see how this model performs for the analysis of the final eBOSS sample. Improvements in the model to account for the shape of the redshift error distribution would also be valuable.

4.3 Observational systematics : study on EZ mocks

This section reviews the weighting scheme applied to the data to treat the potential systematic effects. We denote the total weighting scheme by $\mathcal{W}_{\rm X}$, where the subscript X specifies the different methods to compute the total weight. With this notation, the total weighting scheme used for the DR14 quasar BAO analysis [29] is :

$$\mathcal{W}_{\rm noz} = w_{\rm FKP} \cdot w_{\rm photo} \cdot (w_{\rm cp} + w_{\rm noz} - 1) \tag{4.24}$$

The first term, $w_{\rm FKP} = 1/(1 + \bar{n}(z)P_0)$, is the FKP weight [33] that takes into account the variations of the observed quasar density $n(z)$ across the redshift range and depends on the amplitude of the power spectrum at the scale at which the FKP weights optimize the measurements (here we choose $k = 0.14h$Mpc which is the typical scale at which the BAO signal is well detected which gives $P_0 = 6 \times 10^3 h^{-3}{\rm Mpc}^3$). The second term,$w_{\rm photo}$, is a photometry weight that corrects for the variation of the depth across the survey ; $w_{\rm cp}$, is a weight that accounts for the quasar targets that could not be measured due to fiber collision ; and $w_{\rm noz}$ is a weight that accounts for the confirmed quasars for which a secure redshift could not be determined. We will show that the use of the redshift efficiency, $\epsilon(x, y)$, across the focal plane as a weight, $w_{\rm focal} = 1/\epsilon(x, y)$ is more appropriate than $w_{\rm noz}$. We adopt the following definition of the total weight $\mathcal{W}_{\rm focal}$:

$$\mathcal{W}_{\rm focal} = w_{\rm FKP} \cdot w_{\rm photo} \cdot w_{\rm cp} \cdot w_{\rm focal} \tag{4.25}$$

Each quasar in the DR14 catalog is thus weighted by $\mathcal{W}_{\rm focal}$ to correct for any spurious variation of the quasar densities and to provide a more isotropic selection. For the random catalog, we apply the FKP weight alone as it corresponds to a Poisson sampling which should not be affected by inhomogeneities in the selection. In tests defined throughout the following subsections, using the ability to test against unbiased samples, we will demonstrate that this new weight reduces systematic effects on the quadrupole by a factor of three. Furthermore, the region close to $\mu = 1$ is responsible for the remaining systematic shift, and we propose a method to take this into account. This latest weighting scheme, which corresponds to the one we will adopt for the fitting of the data, is referred as $\mathcal{W}_{{\rm focal}-\mu}$.

4.3.1 Production of EZ mocks with synthetic observational features

In section 2.6.2, we listed some observational artefacts that can bias our measurements, among them the fiber collisions that happen when objects are closer than $62''$ and the spectroscopic completeness when redshifts have not been assigned although spectra have been obtained.

The probability of obtaining a reliable redshift from a spectrum depends on both observational and instrumentation parameters affecting the S/N. When the redshift from an identified quasar cannot be secured, the nearest quasar neighbour is marked such that we can track redshift efficiency and study the weighting scheme to take this into account. Here we extend the treatment applied in previous analyses and search for dependencies with the position in the focal plane. The redshift efficiency or spectroscopic completeness is defined as the ratio between the number of objects with a secured redshift to the number of quasars that received a fiber :

$$\epsilon = \frac{N_{\rm good}}{N(w_{zf}=1) + 2 \cdot N(w_{zf}=2)} \tag{4.26}$$

with $N_{\rm good}$ the number of quasars with robust redshift, and $N(w_{zf} = 1, 2$ the number of quasars without or with a neighbour with a redshift failure. This expression allows for the calculation of the redshift efficiency from the released catalog. The variation of the redshift efficiency for groups of fibers of the spectrographs is displayed in figure 4.10. We observe a decrease in this probability near the side edges of the focal plane. The reason for this behavior is that the light transmitted through fibers near the side-edges of the focal plane arrives near the edges of the CCD, where the optical performance inside the spectrographs is slightly degraded, leading to a larger point spread function and optical aberrations, as already revealed in [34]. It confirms the findings of [11] that the quasar redshift efficiency is lower at the edges of the two spectrographs. Furthermore, the efficiency of the first spectrograph is found to be significantly lower for SGC observations. The variation of redshift efficiency across the focal plane is shown in the bottom panels of Figure 4.10. Regions with lower efficiency are at the left and right sides of the focal plane which correspond to edges of the spectrographs.

To study these systematic effects, we use a more realistic set of EZ and QPM mocks where the plate geometry of the actual survey is applied to retrieve coordinates in the focal plane for each object ensuring that we imprint the same tile distribution of the data in the mocks. From these coordinates, one can determine whether the object belongs to a sector of overlapping regions, in case it does, the particle is randomly assigned to an overlapping plate. The collision-pair effect is applied to particles within $62''$ and which both fall into non-overlapping regions. One particle is removed and the other is up-weighted ($w_{\rm cp}$ weight). Redshift-failure effect is applied by assigning the plate coordinates ($x_{\rm foc}$,$y_{\rm foc}$) to each particle in the mocks and a weight is applied on them by tracking the variation of redshift efficiency across the focal plane shown in the bottom panel of figure 4.10. The above procedure provides the possibility to tag objects in collision and to downsample objects according to the redshift efficiency directly.

4.3.2 Spectroscopic completeness : impact of redshift failures

To study the impact of the spectroscopic completeness we use the special set of EZ mocks that includes the redshift failures. Figure 4.11 shows the difference between the measured correlation function to the correlation function without redshift failures and fiber collisions (both estimated with the EZ mocks). For the quadrupole, using the up-weighting of the nearest neighbor ($\mathcal{W}_{\rm noz}$, red curves) yields a systematic shift of 8% at large scales. An effect is also observed on the monopole but, at first order, it only affects the bias determination. The hexadecapole displays a large effect, although the offset is well within the statistical precision. Results on the fit parameters for the 1,000 EZ mocks are summarized in table 4.5 and exhibit a large shift (e.g. $\Delta f\sigma_8 = 0.105$) for the 3-multipole

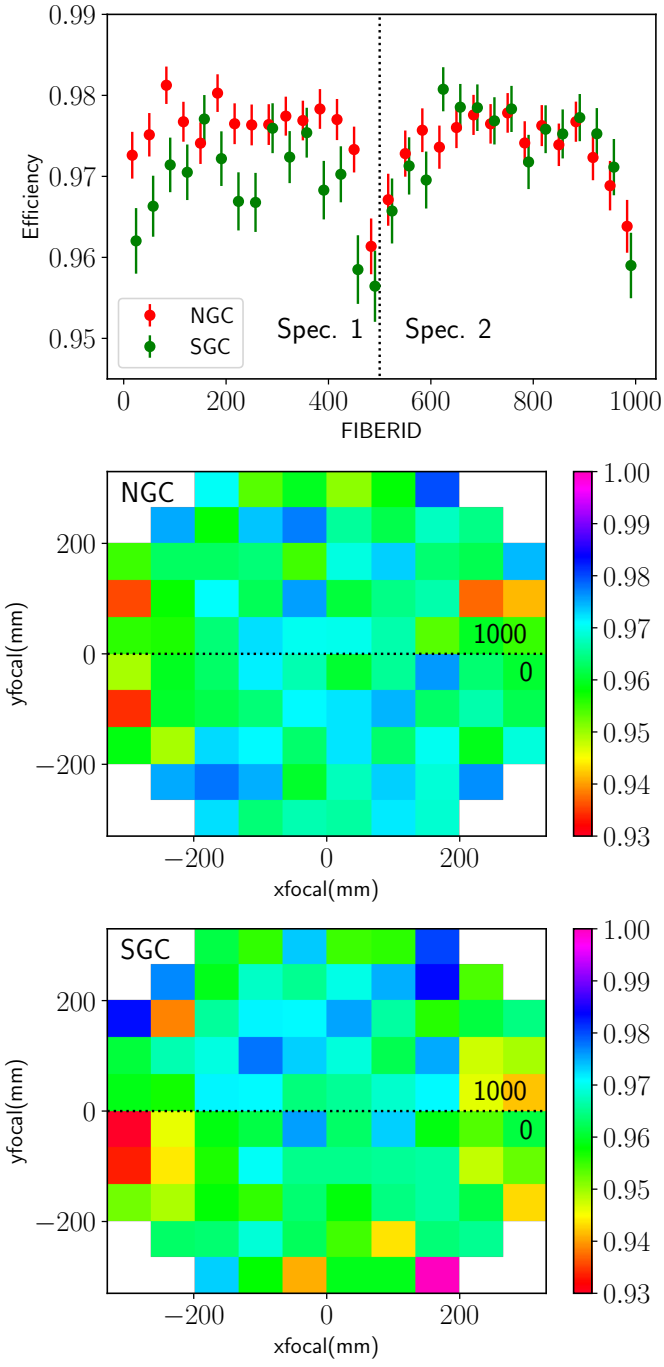


Figure 4.10 – Top panel : Redshift efficiency as a function of the fiber number. The vertical dotted line shows the delimitation between the 2 spectrographs. Bottom panels : Redshift efficiency as a function of the focal plane coordinates for the NGC (middle panel) and SGC (lower panel). The fiber number goes clockwise from 0 to 1000.

case which exceeds even the statistical precision of our measurement. For the 3-wedge analysis the shifts are smaller but still large w.r.t. our precision, especially on $f\sigma_8$.

In the proposed modified weighting scheme the observed quasars are weighted by the inverse of the efficiency calculated from the coordinates of the object in the focal plane. The results are presented ($\mathcal{W}_{\rm focal}$, green curves) in figure 4.11, which reveals a reduction of a factor three of the effect on the quadrupole. As a consequence the average shift estimated from the mocks is decreased to $\Delta f\sigma_8 = 0.033$ (resp. $\Delta f\sigma_8 = 0.013$) for the 3-multipole (resp. 3-wedge) analysis. The parameters $\alpha_\parallel$ and $\alpha_\perp$ are also shifted by 0.02 in the case of the 3-multipole analysis, probably as a consequence of the sensitivity of the hexadecapole to these parameters.

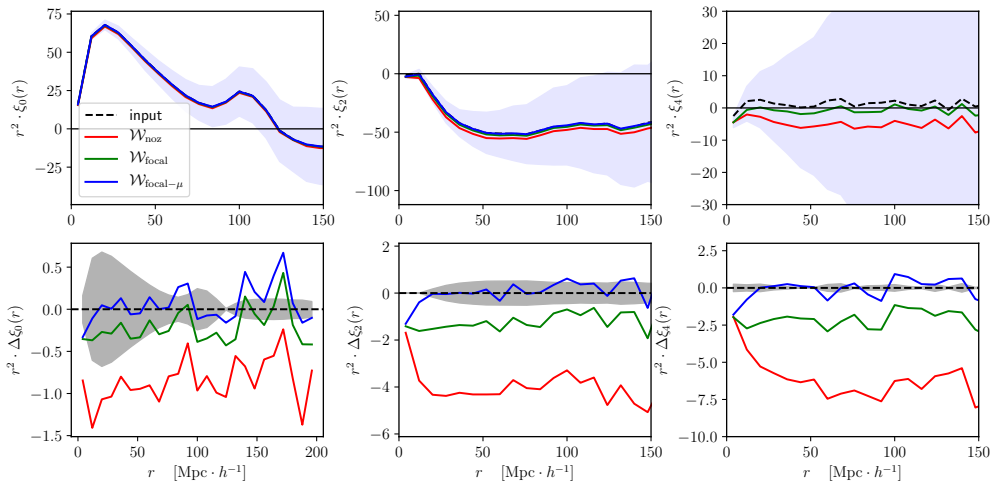


FIGURE 4.11 – Effect of the different weighting schemes on the correlation function multipole (left : monopole, middle : quadrupole, right : hexadecapole). $\mathcal{W}_{\rm noz}$ (red curves) : upweighting of the nearest neighbour for redshift failures. $\mathcal{W}_{\rm focal}$ (green curves) : weight according to the inverse of the spectroscopic efficiency. $\mathcal{W}_{\rm focal-\mu}$ (blue curves) : same as $\mathcal{W}_{\rm focal}$ but the $\mu > 1 - 1/480$ region is removed as described in the text. The light-blue shaded bands on the top plots represent the dispersion of the mocks. Bottom plots : difference between each weighting scheme and the input ; the shaded bands represent a $\pm 1\%$ effect for the monopole and the quadrupole and a $\pm 10\%$ for the hexadecapole.

4.3.3 Finite size of fibers : Impact of close-pairs

In previous analyses, unmeasured targets due to fiber collision are corrected by increasing by one unit the weight of the identified quasar in the collision group. This approach means that any target within 62" of a measured quasar will be displaced along the LOS and brought to the position of the measured quasar. This action inevitably creates a lack of objects at all scales and at $\mu \simeq 1$ and hence will affect the correlation function evaluation. In their measurement of $f\sigma_8$ at small scales where the effect of collisions is large, [35] redefined the correlation function multipoles by excluding the region above a given threshold $\mu_s(s)$ depending on the separation s and defined by the minimum angular distance between two objects ($62''$). Here we adopt a similar approach, but for simplicity we recalculate the value of the correlation function for the last μ-bin in the same manner for

TABLE 4.5 – Effect on the EZ mocks of the different weighting schemes to mitigate systematic effects arising from spectroscopic completeness and fiber collisions. The values and the errors are obtained from 1000 realisations. The reference is given by the same set of mocks but where neither fiber collisions nor spectroscopic completeness are considered.

3-multipole	$f\sigma_8$ ($\Delta f\sigma_8$)	$\alpha_\parallel$ ($\Delta\alpha_\parallel$)	$\alpha_\perp$ ($\Delta\alpha_\perp$)
reference	0.3733±0.0022	0.9950±0.0023	0.9926±0.0020
$\mathcal{W}_{\rm noz}$	+0.1050	-0.0522	0.0559
$\mathcal{W}_{\rm focal}$	+0.0338	-0.0169	+0.0184
$\mathcal{W}_{{\rm focal}-\mu}$	-0.0003	+0.0009	-0.0007
3-wedge	$f\sigma_8$ ($\Delta f\sigma_8$)	$\alpha_\parallel$ ($\Delta\alpha_\parallel$)	$\alpha_\perp$ ($\Delta\alpha_\perp$)
reference	0.3784±0.0031	0.9966±0.0028	0.9963± 0.0025
$\mathcal{W}_{\rm noz}$	0.0424	-0.0086	0.0158
$\mathcal{W}_{\rm focal}$	0.0130	-0.0007	0.0050
$\mathcal{W}_{{\rm focal}-\mu}$	-0.0004	0.0029	-0.0003

TABLE 4.6 – eBOSS DR14 quasar sample : Effect on the data of the different weighting schemes obtained from the 3-multipole and 3-wedge analyses. Differences are calculated w.r.t. the $\mathcal{W}_{{\rm focal}-\mu}$ case and a given in parentheses.

3-multipole	$f\sigma_8$	$\alpha_\parallel$	$\alpha_\perp$
$\mathcal{W}_{\rm noz}$	$0.436^{+0.071}_{-0.072}$ (0.024)	$0.999^{+0.078}_{-0.070}$ (−0.015)	$1.031^{+0.050}_{-0.048}$ (0.006)
$\mathcal{W}_{\rm focal}$	$0.426^{+0.070}_{-0.070}$ (0.014)	$1.014^{+0.070}_{-0.063}$ (0.000)	$1.030^{+0.050}_{-0.048}$ (0.005)
$\mathcal{W}_{{\rm focal}-\mu}$	$0.412^{+0.069}_{-0.070}$ –	$1.014^{+0.070}_{-0.062}$ –	$1.025^{+0.049}_{-0.048}$ –
3-wedge	$f\sigma_8$	$\alpha_\parallel$	$\alpha_\perp$
$\mathcal{W}_{\rm noz}$	$0.343^{+0.084}_{-0.088}$ (−0.021)	$1.089^{+0.141}_{-0.097}$ (0.035)	$1.008^{+0.053}_{-0.053}$ (−0.006)
$\mathcal{W}_{\rm focal}$	$0.365^{+0.082}_{-0.083}$ (0.001)	$1.064^{+0.107}_{-0.081}$ (0.010)	$1.015^{+0.052}_{-0.052}$ (0.001)
$\mathcal{W}_{{\rm focal}-\mu}$	$0.364^{+0.081}_{-0.081}$ –	$1.054^{+0.101}_{-0.078}$ –	$1.014^{+0.052}_{-0.052}$ –

all bins in separation.

At $z \simeq 1.5$, a $62''$ radius exclusion corresponds to 0.54 Mpc ; when considering scales larger than 20 h^{-1}Mpc, we observed that pairs for which $1 - \frac{1}{480} < \mu < 1$ are affected by the upweighting due to close pairs. Since we use 30 μ bins in our analysis, the affected orientation correspond to 1/16 of the last μ bin. To mitigate this effect, we discard the paircounts in this region and rescale the counts of the last μ bin by 16/15.

The results obtained after this correction was applied to the EZ mocks are shown as the blue curves of figure 4.11. With this method, for scales larger than 15 h^{-1}Mpc, the true quadrupole is recovered to an accuracy better than 1%, and no systematic behaviour is found on the monopole. The result on the cosmological parameters extracted from the fit of the 1000 EZ mocks with our model are in agreement with the reference with $\Delta f\sigma_8 < 0.001$, $\Delta\alpha_\parallel < 0.003$ and $\Delta\alpha_\perp < 0.001$. This method allows for a mitigation of the effect of fiber collisions and redshift efficiency variations across the focal plane at the level where it will not be a limitation even when the full eBOSS quasar sample will be available.

Thanks to the improved treatment of redshift failures and close pairs, we report no observational systematics for the analysis in configuration space presented in [1]. The systematic budget thus comes from the modelling systematics only which represents about half of the statistical precision.

4.4 Study of systematics on data

4.4.1 Weighting scheme

The different weighting schemes are also applied on the data and the fits results are given in table 4.6. The differences in the fits between the weighting schemes are found to be smaller than in the mocks. The largest differences between favored schemes $\mathcal{W}_{\text{focal}-\mu}$ and $\mathcal{W}_{\text{focal}}$ are $\Delta f\sigma_8 = 0.014$ and $\Delta\alpha_{\parallel} = 0.010$; these differences represent only 20% of the statistical precision. From the distribution of differences in the mocks what is observed in the data is not unusual, although an alternative explanation is that the effect of close pairs and redshift failures is somehow magnified in our improved set of EZ mocks. In the following we use $\mathcal{W}_{\text{focal}-\mu}$ weighting scheme as our reference. For consistency with other analyses which do not employ this weighting scheme we will also present the results for the case $\mathcal{W}_{\text{focal}}$.

4.4.2 Imaging : impact of photometric weights

The impact of the inhomogeneity of the quasar target selection on the observed eBOSS quasar density was first studied by [11] using the first year of eBOSS observations. Following the approach of [36, 37, 38] for BOSS analyses, they introduced a photometric weight w_{photo} according to the 5-σ detection in magnitude for a point source, also called depth. For a point-like SDSS source, its mean value for the g-band (resp. r-band) is $g = 23.1$ (reso. $r = 22.7$). Variations in the quality of the SDSS photometry yields angular variations of the depth that can reach ± 0.8 magnitudes, so some sources can be at the limit of the 5-σ detection. For each filter, the depth can be defined as follows :

$$m_{5\sigma} = -2.5\log(AS\sqrt{\Phi_{\text{sky}}}10^{0.4kAirmass}) - m_{\text{ext}} \tag{4.27}$$

where S is the full width at half maximum of the point-source function (also called *seeing*), Φ_{sky} is the observed sky flux without point-like sources, $Airmass$ is the air density column and $-m_{\text{ext}}$ is the correction that needs to be applied on magnitudes from absorption and dust due to the Galactic extinction. The quantities A and k depend on the filter u,g,r,i,z. [11] showed that quasars are more securely identified where the depth is high and Galactic extinction is the variable that has been found to be the most sensitive to differences in depth between the NGC and SGC, so the photometric weights for the DR14 sample have been computed with separate correction for each Galactic cap. By studying the variation of the observed quasar density as a function of depth which contains the dependence on airmass, seeing and Galactic extinction, one can compute photometric weights based on linear fits according to the dependency with the depth. These weights actually mitigate the systematic errors in the evaluation of the correlation function induced by the variation of the depth across the footprint. It corresponds to the grey squares in figure 4.12 which represents the number density of the DR14 quasar sample after correction as a function of various observational systematics. The dashed red curve is the number density of the DR14 quasar sample before correction, we can see that weighting for depth and Galactic extinction removes correlations with other potential systematic quantities.

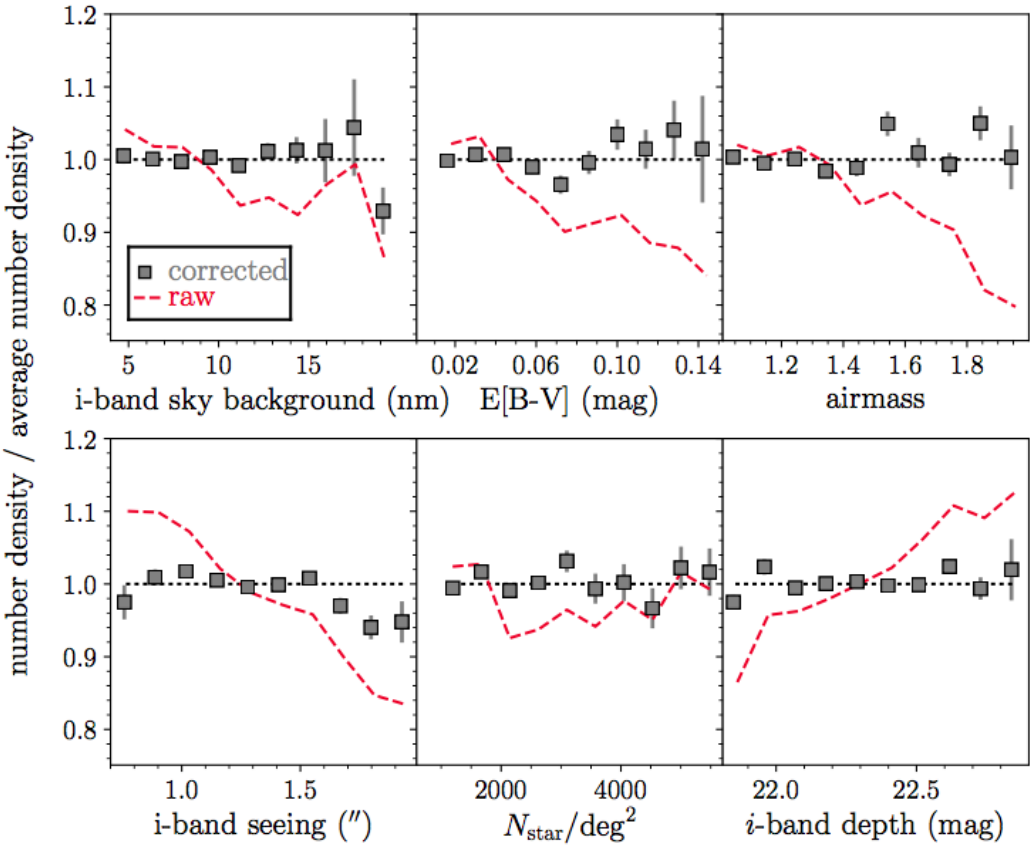


FIGURE 4.12 – Relationship between the number density of the DR14 quasar sample and various observational systematics before (dashed red curves) and after (grey squares) correction applied on the depth and the Galactic extinction. The results that are shown are a combination of the NGC and SGC.

We first determine the dependency with depth and then with Galactic extinction, after applying the weights for depth. The total photometric weight is the multiplication of these two effects :

$$w_{\text{photo}} = \frac{1}{(A_d + dB_d)(A_e + eB_e)} \tag{4.28}$$

where d is the g-band depth and e is the Galactic extinction. The best-fit coefficients are :

cap	A_d	B_d	A_e	B_e
NGC	-3.52	0.195	1.045	-2.01
SGC	-6.20	0.310	1.052	-1.00

The impact of the photometric weights on the clustering statistics is shown in figure 4.13. The top panel represents the distribution of photometric systematic weights computed for both Galactic caps, and for the two regions of the SGC separately ; the spread of weights is much larger for the SGC. By computing the photometric weights in each cap separetely, we can correct for the variation in targeting efficiency due to differences in imaging properties. As explained in [39] due to better observing conditions, we expect the target selection to be more efficient in the NGC. The bottom panels of figure 4.13 show the impact of the photometric weights on the correlation function for the NGC (left panel) and SGC (right panel). The effect of weights on the correlation function is almost constant across the range of separation considered for this analysis. The effect on the correlation function for wedges in μ is similar for all wedges ; as a consequence, the effects on the quadrupole and on the hexadecapole are small although some spread is observed in the correction for the SGC. As observed in [11], the effect on the monopole is much larger for the SGC than for the NGC, but the corrected correlation function shows no remaining systematics within the current precision (e.g., the top panel of Figure 5.1 in the next section).

Additional tests were conducted on the WISE photometry which also enters the target selection algorithm. We used the method developed in [40] to estimate the weights from the linear regression of the target density w.r.t. the photometric parameters including WISE, and no significant effect was observed. Moreover, the regions where there is some contamination from the moon (mostly in the SGC) were removed ; this deletion produced no impact on our results.

4.4.3 Redshift estimates

As explained in section 2.5, we can use three redshift estimates to measure the clustering of the DR14 quasar sample. We adopt the redshift z as the reference throughout this analysis and compare its results with catalogs where the redshift is taken to be z_{MgII} (resp. z_{PCA}) whenever it is available (i.e. 80% of the time) and z otherwise such that these catalogs have the same objects.

The results in table 4.7 are consistent within 1σ, although the results from z_{PCA} exhibit a stronger deviation than z_{MgII}. This behaviour could be an argument in favor of the astrophysical motivations to use MgII-based redshift, since it is supposed to be the more systematics-free redshift estimate, but further investigation on the reliability of the MgII line across our entire redshift range is required before stating firm conclusions. In addition, these measurements that use redshift estimates should not be considered as independent, and because we lack equivalent different redshift estimates for mocks we cannot simply combine the redshifts.

Differences in clustering between z_{MgII} (resp. z_{PCA}) w.r.t z can be compared to the dispersion due to different realizations of the same mock for a given distribution of redshift

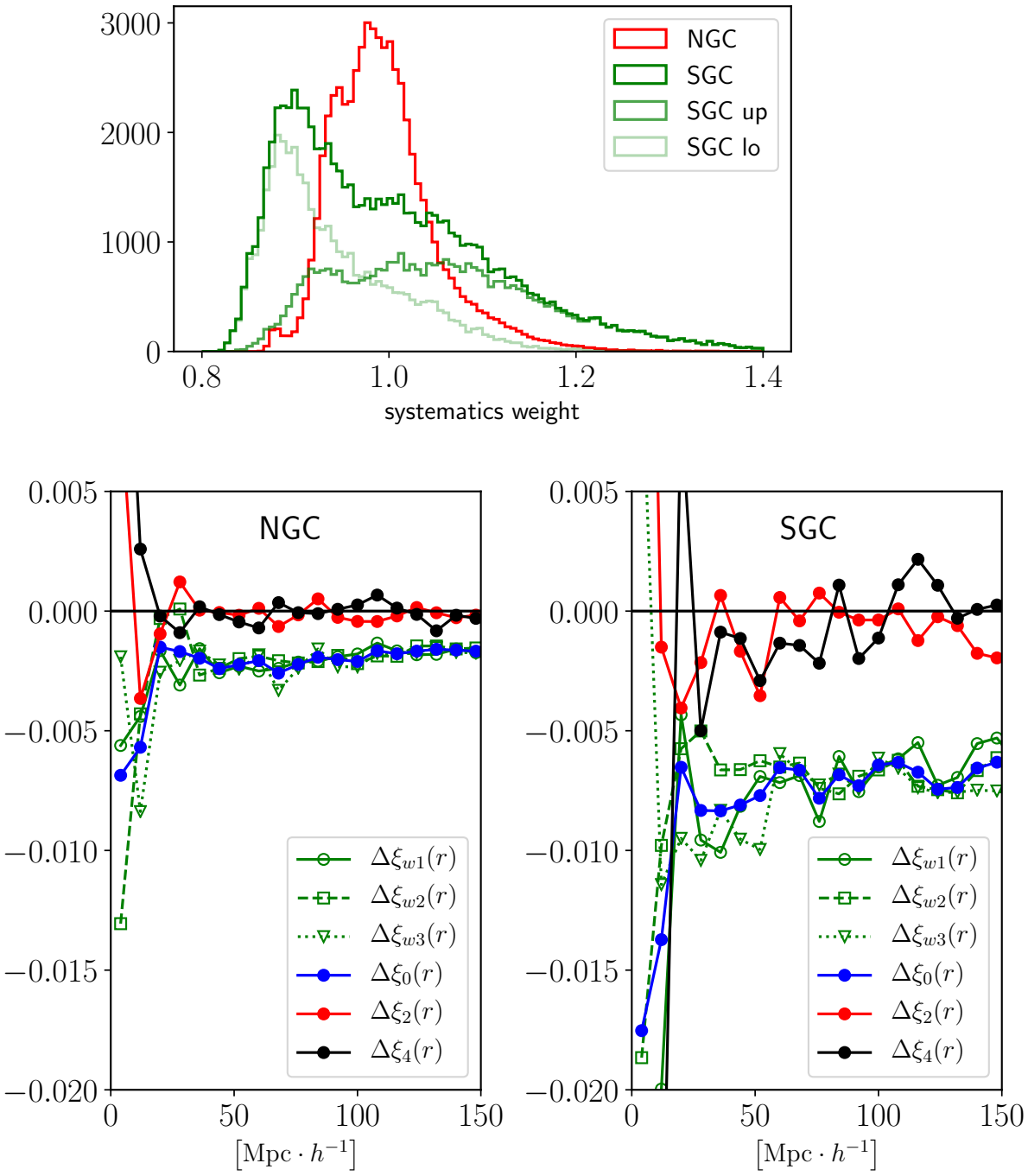


Figure 4.13 – Top panel : Distribution of the photometric weights applied to the data to correct for inhomogeneity in the depth of the photometric sample used at the targeting stage. Bottom panels : effect of photometric weights on the monopole, quadrupole and wedges of the correlation function for the NGC (left) and SGC (right). Note that the correlation function is not multiplied by s on these plots.

uncertainties. For a specific OuterRim mock catalog, we can draw several realizations of a given redshift smearing on the same mock. We investgate the case with 13% satellite fraction and with a Gaussian SRD distribution for redshift uncertainties since it is the closest configuration to the data and we draw 30 different realizations of the same smearing. We then examine the dispersion on the Legendre multipoles $\ell = 0, 2, 4$ and wedges which corresponds to the grey envelope in figure 4.14 for the three multipoles and in figure 4.15 for the three wedges. The differences in clustering using different redshift estimates lie within the dispersion expected from statistically independent redshift smearing and they do not show any systematic trend. We conclude that differences between the results of the fit with the different redshift estimates are due to statistics and we do not quote an additional systematic uncertainty.

4.4.4 Additional tests

Table 4.7 summarizes the different tests we perform on data to compare to the reference and to study the robustness of our measurements. In particular, we review at the following effects :

— **Isotropic analysis** : As a consistency check, using $\alpha_{\rm iso} \simeq \alpha_{\parallel}^{1/3}\alpha_{\perp}^{2/3}$ with the reference values from the anisotropic fitting of the three multipoles for instance for $\alpha_{\parallel}$ and $\alpha_{\perp}$, we compute $\alpha_{\rm iso} = 1.021 \pm 0.057$, which matches well the result from the isotropic fit. The effect on $f\sigma_8$ is also consistent, and no significant shift is reported.

— **Fixing the fiducial cosmology** produces consistent results with the anisotropic and isotropic cases, and as expected given the degeneracy between the AP parameters and $f\sigma_8$, this approach provides a better constraint on $f\sigma_8$. However, if one wishes to constrain modified gravity models based on different assumptions than the one of ΛCDM-GR for structure formation, one must use the results obtained by the full anisotropic clustering using AP parameters.

— **Effect of covariance matrix** : There is no significant effect on the cosmological parameters $f\sigma_8$, $\alpha_{\parallel}$ and $\alpha_{\perp}$ when using the covariance matrix from the 400 QPM mocks instead of the 1,000 EZ mocks.

— **Effect of redshift resolution** : When fixing $\sigma_{\rm tot}$ to the best-fitting values, the precision on $\alpha_{\parallel}$ is improved by 30% as seen in tests on the Outer Rim catalogs. This results provides clear motivation to improve our knowledge of the redshift uncertainty for future quasar samples.

— **Effect of F'' prescription** : as shown in the model, there is no significant difference on the fitted cosmological parameters when using PS mass function instead of ST. We do not report any result when letting F'' free because we are not sufficiently sensitive to this parameter to derive useful constraints. In addition, since F'' accounts for non-linearities in the bias model at small scales, it may be degenerate with $\sigma_{\rm tot}$.

The fits are also performed on the two Galactic caps separately whose results are given at the end of table 4.7 for the 3-multipole and 3-wedge fits. The fit parameters are in agreement, although the χ^2 of the fit of the SGC using 3-multipole reaches $\chi^2 = 55.0(\text{d.o.f} = 40)$. We conducted extensive tests in order to isolate a potential source for this effect. This increase in χ^2 has been located in the $\delta > 10$ deg area of the SGC which is the region where the spread of the photometric weights distribution is the greatest. After removing regions of extreme values of the systematic weights, with moon contamination in WISE photometry, or regions of high Galactic extinction, no obvious source could be identified so we kept those regions in the analysis. Figure 4.16 compares the χ^2 on the data in each cap (dashed

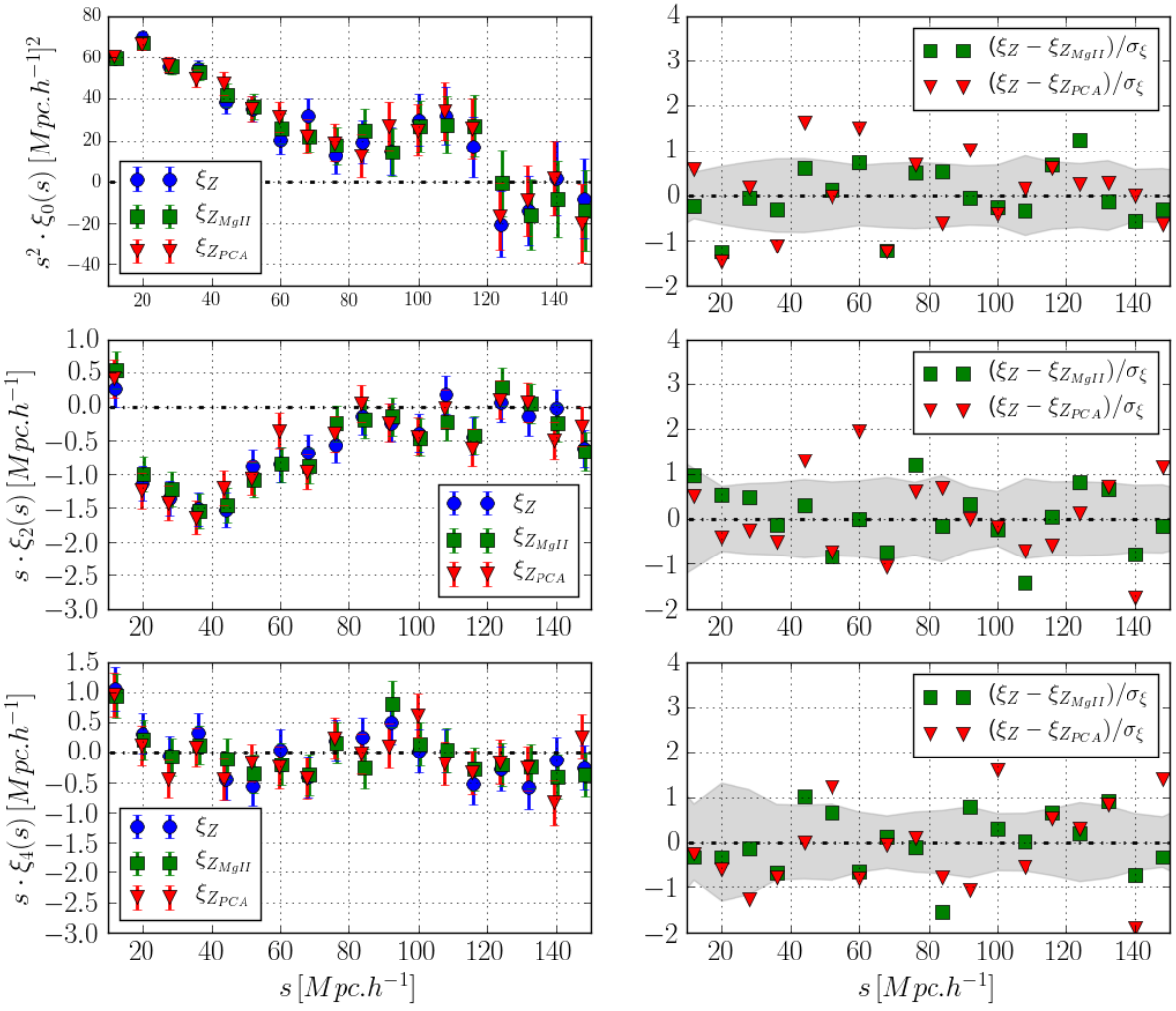


FIGURE 4.14 – Right panel : Monopole (top), quadrupole (middle) and hexadecapole (bottom) for z (blue), $z_{\rm MgII}$ (green) and $z_{\rm PCA}$ (red). Left panel : Difference $\xi_{z_{\rm MgII}} - \xi_z$ and $\xi_{z_{\rm PCA}} - \xi_z$ divided by the error using EZ mocks, compared to the dispersion of 30 realizations for the same mock with a Gaussian redshift uncertainties distribution according to to the SRD for the Legendre multipoles $\ell = 0, 2, 4$. The differences in clustering are consistent with the expected dispersion from statistically independent redshift smearing.

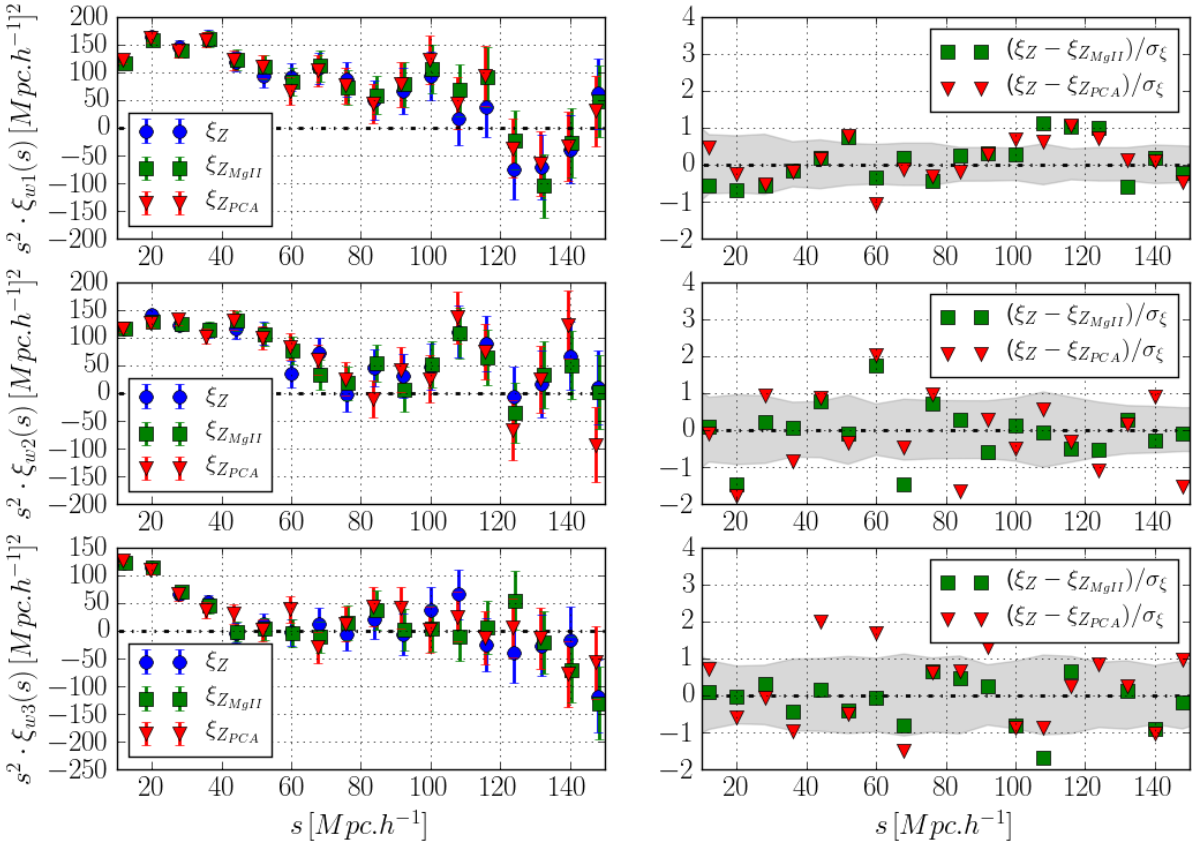


FIGURE 4.15 – Right panel : Wedge $0 < \mu < 1/3$ (top), wedge $1/3 < \mu < 2/3$ (middle) and wedge $2/3 < \mu < 1$ (bottom) for z (blue), $z_{\rm MgII}$ (green) and $z_{\rm PCA}$ (red). Left panel : Difference $\xi_{z_{\rm MgII}} - \xi_z$ and $\xi_{z_{\rm PCA}} - \xi_z$ divided by the error using EZ mocks, compared to the dispersion of 30 realizations for the same mock with a Gaussian redshift uncertainties distribution according to to the SRD for the three wedges. The differences in clustering are consistent with the expected dispersion from statistically independent redshift smearing.

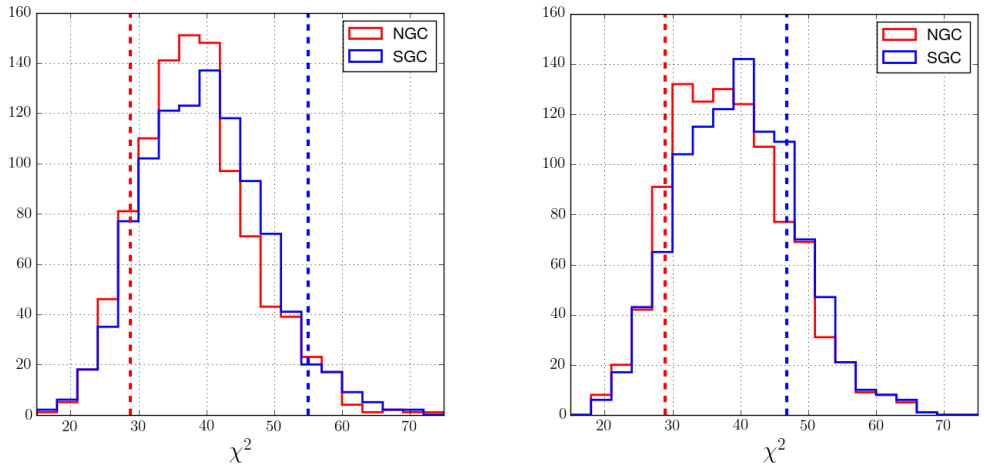


FIGURE 4.16 – Distribrution of the χ^2(d.o.f = 40) of the fits of the 1000 EZ mocks per Galactic cap (solid line) and comparison with the χ^2 obtained from the data (dashed line). Left : for 3-multipole. Right : for 3-wedge. The χ^2 on the data are found to be within the distribution of the EZ mocks, even for the χ^2 in the SGC (blue dashed line) which is larger for the 3-multipole analysis. Results on each cap thus represent a statistical realization of the EZ mocks.

line) with the χ^2 distribution obtained for the results of the 1,000 EZ mocks (solid) by cap (NGC in red and SGC in blue) for the 3-multipole (top panel) and 3-wedge (bottom panel) fits. As described in Section 4.1.2, the NGC and SGC EZ mocks are created from separate simulations whose bias parameters have been adjusted on the observed DR14 eBOSS quasar clustering on each cap directly. It is clearly visible in figure 4.16 that the χ^2 in the SGC (blue dashed) for the 3-multipole analysis is large but not unusual compared to the EZ mocks distribution.

Summary

We demonstrate the applicability of the CLPT-GS model for dark matter halos of masses of the order of $10^{12.5}M_\odot$ hosting eBOSS quasar tracers at mean redshift $z \simeq 1.5$. In order to estimate the systematic error budget related to the RSD modeling, we use the N-body Outer Rim simulation to test the predictions of CLPT in real space and then evaluate the performance of the model in redshift space using a hundred mock catalogs created for that purpose. We also propose to move beyond the traditional weighting scheme that was used for BOSS galaxies and the BAO measurement with the DR14 sample to account for redshift failures and close-pairs. We validate the procedure on a thousand EZ mock catalogs. This approach allows the observational systematics to be much smaller than the current statistical precision. In the light of the study of potential systematics performed on both synthetic mock catalogs and the data sample, all the tests provide compatible results, suggesting that none of the options we tested affects our estimate of the uncertainty on our cosmological parameters or bias our results by more than 1σ, within statistics.

Bibliographie

[1] P. Zarrouk, E. Burtin, H. Gil-Marín, A. J. Ross, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : measure-*

TABLE 4.7 – Results of the fit of the data when changing cosmological assumptions, redshifts estimates, covariance matrix determination, and second order bias. The upper part (resp. lower part) of the table presents results for the 3-multipole (resp. 3-wedge) analysis. Results for each individual Galactic cap are given at the end of each table.

config	$b\sigma_8$	$f\sigma_8$	$\alpha_\parallel$	$\alpha_\perp$	$\sigma_{\rm tot}$	χ^2/dof
3-multipole						
NGC+SGC ref	$1.042^{+0.059}_{-0.056}$	$0.412^{+0.069}_{-0.070}$	$1.014^{+0.070}_{-0.062}$	$1.025^{+0.049}_{-0.048}$	$6.00^{+1.19}_{-1.41}$	$40.5(45-5)$
NGC+SGC isotropic	$1.044^{+0.056}_{-0.053}$	$0.406^{+0.054}_{-0.053}$	$\alpha_{\rm iso} = 1.021^{+0.039}_{-0.037}$	–	$6.10^{+0.89}_{-1.12}$	$40.5(45-4)$
NGC+SGC fiducial	$1.019^{+0.030}_{-0.030}$	$0.398^{+0.050}_{-0.051}$	fixed	fixed	$5.91^{+0.81}_{-1.09}$	$40.9(45-3)$
NGC+SGC covQPM	$1.054^{+0.068}_{-0.060}$	$0.386^{+0.069}_{-0.071}$	$1.055^{+0.083}_{-0.068}$	$1.022^{+0.051}_{-0.049}$	$6.60^{+1.21}_{-1.35}$	39.9
NGC+SGC '$z_{\rm PCA}$'	$0.997^{+0.066}_{-0.065}$	$0.387^{+0.072}_{-0.073}$	$0.988^{+0.085}_{-0.080}$	$1.005^{+0.048}_{-0.048}$	$5.37^{+1.36}_{-1.74}$	39.3
NGC+SGC '$z_{\rm MgII}$'	$0.966^{+0.074}_{-0.066}$	$0.424^{+0.070}_{-0.073}$	$0.972^{+0.095}_{-0.079}$	$0.994^{+0.052}_{-0.049}$	$6.27^{+1.33}_{-1.38}$	31.8
NGC+SGC $\sigma_{\rm tot}$ fixed	$1.042^{+0.057}_{-0.055}$	$0.412^{+0.069}_{-0.070}$	$1.014^{+0.051}_{-0.045}$	$1.025^{+0.049}_{-0.047}$	6.00 (fixed)	40.5
NGC+SGC F''_{PS}	$1.042^{+0.059}_{-0.056}$	$0.412^{+0.069}_{-0.069}$	$1.015^{+0.070}_{-0.062}$	$1.025^{+0.049}_{-0.047}$	$6.06^{+1.18}_{-1.39}$	40.5
NGC AP	$0.960^{+0.083}_{-0.076}$	$0.440^{+0.083}_{-0.084}$	$0.950^{+0.102}_{-0.083}$	$0.992^{+0.058}_{-0.054}$	$6.30^{+1.50}_{-1.54}$	28.8
SGC AP	$1.142^{+0.085}_{-0.078}$	$0.383^{+0.097}_{-0.096}$	$1.048^{+0.100}_{-0.077}$	$1.086^{+0.071}_{-0.072}$	$3.68^{+2.61}_{-3.68}$	55.0
3-wedge						
NGC+SGC ref	$1.069^{+0.066}_{-0.059}$	$0.364^{+0.081}_{-0.081}$	$1.054^{+0.101}_{-0.078}$	$1.014^{+0.052}_{-0.052}$	$6.08^{+1.57}_{-1.74}$	$37.8(45-5)$
NGC+SGC isotropic	$1.060^{+0.055}_{-0.054}$	$0.385^{+0.057}_{-0.056}$	$\alpha_{\rm iso} = 1.027^{+0.039}_{-0.037}$	–	$5.70^{+1.07}_{-1.42}$	$38.0(45-4)$
NGC+SGC fiducial	$1.028^{+0.031}_{-0.031}$	$0.374^{+0.053}_{-0.054}$	fixed	fixed	$5.42^{+1.00}_{-1.42}$	$38.5(45-3)$
NGC+SGC covQPM	$1.042^{+0.080}_{-0.063}$	$0.383^{+0.078}_{-0.078}$	$1.029^{+0.119}_{-0.078}$	$1.027^{+0.056}_{-0.055}$	$6.01^{+1.75}_{-1.84}$	42.4
NGC+SGC '$z_{\rm PCA}$'	$1.064^{+0.067}_{-0.075}$	$0.277^{+0.097}_{-0.092}$	$1.130^{+0.113}_{-0.123}$	$0.963^{+0.054}_{-0.055}$	$6.85^{+1.65}_{-1.96}$	44.1
NGC+SGC '$z_{\rm MgII}$'	$1.027^{+0.097}_{-0.115}$	$0.362^{+0.124}_{-0.113}$	$1.094^{+0.168}_{-0.181}$	$0.975^{+0.059}_{-0.055}$	$7.71^{+2.07}_{-2.50}$	38.6
NGC+SGC $\sigma_{\rm tot}$ fixed	$1.069^{+0.062}_{-0.058}$	$0.364^{+0.081}_{-0.080}$	$1.054^{+0.066}_{-0.056}$	$1.014^{+0.050}_{-0.050}$	6.09 (fixed)	37.8
NGC+SGC F''_{PS}	$1.070^{+0.066}_{-0.059}$	$0.364^{+0.081}_{-0.081}$	$1.055^{+0.101}_{-0.078}$	$1.014^{+0.052}_{-0.052}$	$6.13^{+1.56}_{-1.71}$	37.8
NGC AP	$1.001^{+0.146}_{-0.090}$	$0.373^{+0.108}_{-0.133}$	$1.007^{+0.249}_{-0.122}$	$0.973^{+0.061}_{-0.059}$	$6.63^{+2.99}_{-2.12}$	28.9
SGC AP	$1.143^{+0.087}_{-0.081}$	$0.387^{+0.120}_{-0.110}$	$1.032^{+0.109}_{-0.081}$	$1.094^{+0.085}_{-0.081}$	$3.14^{+3.08}_{-3.14}$	46.8

ment of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2, **477**, 1639 (2018), doi:10.1093/mnras/sty506.

[2] D. Alonso, *CUTE : Correlation Utilities and Two-point Estimation*, Astrophysics Source Code Library (2015), ascl:1505.016.

[3] A. J. S. Hamilton, in *The Evolving Universe*, édité par D. Hamilton (1998), vol. 231 de *Astrophysics and Space Science Library*, p. 185, astro-ph/9708102.

[4] E. A. Kazin, A. G. Sánchez, et M. R. Blanton, *Improving measurements of H(z) and D_A (z) by analysing clustering anisotropies*, **419**, 3223 (2012), doi:10.1111/j.1365-2966.2011.19962.x, arXiv:1105.2037.

[5] M. White, J. L. Tinker, et C. K. McBride, *Mock galaxy catalogues using the quick particle mesh method*, **437**, 2594 (2014), doi:10.1093/mnras/stt2071, arXiv:1309.5532.

[6] J. L. Tinker, E. S. Sheldon, R. H. Wechsler, M. R. Becker, *et al.*, *Cosmological Constraints from Galaxy Clustering and the Mass-to-number Ratio of Galaxy Clusters*, **745**, 16 (2012), doi:10.1088/0004-637X/745/1/16, arXiv:1104.1635.

[7] J. Carlson et M. White, *Embedding Realistic Surveys in Simulations Through Volume Remapping*, **190**, 311 (2010), doi:10.1088/0067-0049/190/2/311, arXiv:1003.3178.

[8] K. S. Dawson, J.-P. Kneib, W. J. Percival, S. Alam, *et al.*, *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey : Overview and Early Data*, **151**, 44 (2016), doi:10.3847/0004-6256/151/2/44, arXiv:1508.04473.

[9] C.-H. Chuang, F.-S. Kitaura, F. Prada, C. Zhao, *et al.*, *EZmocks : extending the Zel'dovich approximation to generate mock galaxy catalogues with accurate clustering statistics*, **446**, 2621 (2015), doi:10.1093/mnras/stu2301, arXiv:1409.1124.

[10] S. M. Croom, B. J. Boyle, T. Shanks, R. J. Smith, *et al.*, *The 2dF QSO Redshift Survey - XIV. Structure and evolution from the two-point correlation function*, **356**, 415 (2005), doi:10.1111/j.1365-2966.2004.08379.x, astro-ph/0409314.

[11] P. Laurent, S. Eftekharzadeh, J.-M. Le Goff, A. Myers, *et al.*, *Clustering of quasars in SDSS-IV eBOSS : study of potential systematics and bias determination*, **7**, 017 (2017), doi:10.1088/1475-7516/2017/07/017, arXiv:1705.04718.

[12] F.-S. Kitaura, S. Rodríguez-Torres, C.-H. Chuang, C. Zhao, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : mock galaxy catalogues for the BOSS Final Data Release*, **456**, 4156 (2016), doi:10.1093/mnras/stv2826, arXiv:1509.06400.

[13] A. J. Ross, W. J. Percival, et M. Manera, *The information content of anisotropic Baryon Acoustic Oscillation scale measurements*, **451**, 1331 (2015), doi:10.1093/mnras/stv966, arXiv:1501.05571.

[14] H. Gil-Marín, J. Guy, P. Zarrouk, E. Burtin, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : structure growth rate measurement from the anisotropic quasar power spectrum in the redshift range 0.8 z 2.2*, **477**, 1604 (2018), doi:10.1093/mnras/sty453, arXiv:1801.02689.

[15] J. Hartlap, P. Simon, et P. Schneider, *Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix*, **464**, 399 (2007), doi:10.1051/0004-6361:20066170, astro-ph/0608064.

[16] W. J. Percival, A. J. Ross, A. G. Sánchez, L. Samushia, *et al.*, *The clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : including covariance matrix errors*, **439**, 2531 (2014), doi:10.1093/mnras/stu112, arXiv:1312.4841.

[17] D. Foreman-Mackey, D. W. Hogg, D. Lang, et J. Goodman, *emcee : The MCMC Hammer*, **125**, 306 (2013), doi:10.1086/670067, arXiv:1202.3665 [astro-ph.IM].

[18] B. A. Reid et M. White, *Towards an accurate model of the redshift-space clustering of haloes in the quasi-linear regime*, **417**, 1913 (2011), doi:10.1111/j.1365-2966.2011.19379.x, arXiv:1105.4165.

[19] J. Carlson, B. Reid, et M. White, *Convolution Lagrangian perturbation theory for biased tracers*, **429**, 1674 (2013), doi:10.1093/mnras/sts457, arXiv:1209.0780.

[20] L. Wang, B. Reid, et M. White, *An analytic model for redshift-space distortions*, **437**, 588 (2014), doi:10.1093/mnras/stt1916, arXiv:1306.1804.

[21] M. White, M. Blanton, A. Bolton, D. Schlegel, *et al.*, *The Clustering of Massive Galaxies at z ˜ 0.5 from the First Semester of BOSS Data*, **728**, 126 (2011), doi:10.1088/0004-637X/728/2/126, arXiv:1010.4915.

[22] S. Alam, F. D. Albareti, C. Allende Prieto, F. Anders, *et al.*, *The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey : Final Data from SDSS-III*, ArXiv e-prints (2015), arXiv:1501.00963.

[23] S. Satpathy, S. Alam, S. Ho, M. White, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : on the measurement of growth rate using galaxy correlation functions*, **469**, 1369 (2017), doi:10.1093/mnras/stx883, arXiv:1607.03148.

[24] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, *et al.*, *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations : Cosmological Interpretation*, **192**, 18 (2011), doi:10.1088/0067-0049/192/2/18, arXiv:1001.4538 [astro-ph.CO].

[25] A. D. Ludlow, J. F. Navarro, R. E. Angulo, M. Boylan-Kolchin, *et al.*, *The mass-concentration-redshift relation of cold dark matter haloes*, **441**, 378 (2014), doi:10.1093/mnras/stu483, arXiv:1312.0945.

[26] J. F. Navarro, C. S. Frenk, et S. D. M. White, *The Structure of Cold Dark Matter Halos*, **462**, 563 (1996), doi:10.1086/177173, astro-ph/9508025.

[27] S. de la Torre et L. Guzzo, *Modelling non-linear redshift-space distortions in the galaxy clustering pattern : systematic errors on the growth rate parameter*, **427**, 327 (2012), doi:10.1111/j.1365-2966.2012.21824.x, arXiv:1202.5559.

[28] J. C. Jackson, *A critique of Rees's theory of primordial gravitational radiation*, **156**, 1P (1972), doi:10.1093/mnras/156.1.1P, arXiv:0810.3908.

[29] M. Ata, F. Baumgarten, J. Bautista, F. Beutler, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : First measurement of Baryon Acoustic Oscillations between redshift 0.8 and 2.2*, ArXiv e-prints (2017), arXiv:1705.06373.

[30] S. Eftekharzadeh, A. D. Myers, M. White, D. H. Weinberg, *et al.*, *Clustering of intermediate redshift quasars using the final SDSS III-BOSS sample*, ArXiv e-prints (2015), arXiv:1507.08380.

[31] B. A. Reid, L. Samushia, M. White, W. J. Percival, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : measurements of the growth of structure and expansion rate at $z = 0.57$ from anisotropic clustering*, **426**, 2719 (2012), doi:10.1111/j.1365-2966.2012.21779.x, arXiv:1203.6641.

[32] Z. Vlah, E. Castorina, et M. White, *The Gaussian streaming model and convolution Lagrangian effective field theory*, **12**, 007 (2016), doi:10.1088/1475-7516/2016/12/007, arXiv:1609.02908.

[33] H. A. Feldman, N. Kaiser, et J. A. Peacock, *Power-spectrum analysis of three-dimensional redshift surveys*, **426**, 23 (1994), doi:10.1086/174036, astro-ph/9304022.

[34] S. A. Smee, J. E. Gunn, A. Uomoto, N. Roe, *et al.*, *The Multi-object, Fiber-fed Spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey*, **146**, 32 (2013), doi:10.1088/0004-6256/146/2/32, arXiv:1208.2233.

[35] B. A. Reid, H.-J. Seo, A. Leauthaud, J. L. Tinker, *et al.*, *A 2.5 per cent measurement of the growth rate from small-scale redshift space clustering of SDSS-III CMASS galaxies*, **444**, 476 (2014), doi:10.1093/mnras/stu1391, arXiv:1404.3742.

[36] A. J. Ross, S. Ho, A. J. Cuesta, R. Tojeiro, *et al.*, *Ameliorating systematic uncertainties in the angular clustering of galaxies : a study using the SDSS-III*, **417**, 1350 (2011), doi:10.1111/j.1365-2966.2011.19351.x, arXiv:1105.2320.

[37] A. J. Ross, W. J. Percival, A. G. Sánchez, L. Samushia, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : analysis of potential systematics*, **424**, 564 (2012), doi:10.1111/j.1365-2966.2012.21235.x, arXiv:1203.6499.

[38] A. J. Ross, F. Beutler, C.-H. Chuang, M. Pellejero-Ibanez, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : observational systematics and baryon acoustic oscillations in the correlation function*, **464**, 1168 (2017), doi:10.1093/mnras/stw2372, arXiv:1607.03145.

[39] A. D. Myers, N. Palanque-Delabrouille, A. Prakash, I. Pâris, *et al.*, *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey : Quasar Target Selection*, **221**, 27 (2015), doi:10.1088/0067-0049/221/2/27, arXiv:1508.04472.

[40] A. Prakash, T. C. Licquia, J. A. Newman, A. J. Ross, *et al.*, *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey : Luminous Red Galaxy Target Selection*, **224**, 34 (2016), doi:10.3847/0067-0049/224/2/34, arXiv:1508.04478.

Chapitre 5

Clustering measurements of the eBOSS DR14 quasar sample

Contents

In this chapter, section 5.1 presents the main results of this thesis work, which have been published in [1], and their consistency in configuration space between different methods using the same RSD model and fitting conditions. Then, section 5.2 examines the consistency between this work and the other analyses using the eBOSS DR14 quasar sample, both for anisotropic measurements from the full-shape analysis and the isotropic BAO

Table 5.1 – Results for the anisotropic full-shape analysis

type	config	$b\sigma_8$	$f\sigma_8$	$\alpha_{\rm par}$	$\alpha_{\rm perp}$	$\sigma_{\rm tot}$	χ^2/dof
3-multipole	NGC+SGC	$1.038^{+0.060}_{-0.057}$	$0.426^{+0.070}_{-0.070}$	$1.012^{+0.071}_{-0.064}$	$1.031^{+0.050}_{-0.048}$	$5.94^{+1.19}_{-1.40}$	$42.9/(45-5)$
3-wedge	NGC+SGC	$1.068^{+0.066}_{-0.062}$	$0.363^{+0.082}_{-0.081}$	$1.054^{+0.102}_{-0.078}$	$1.013^{+0.052}_{-0.052}$	$6.10^{+1.57}_{-1.73}$	$37.5/(45-5)$

measurement presented in [2]. In section 5.3 we propose to combine our measurements in configuration space with our companion paper in Fourier space [3] as an illustrative example of the methodology to combine results from the same data sample and of the gain we can expect on the achievable precision for cosmological parameters. Eventualy, in section 5.4, we compare our configuration space measurements of the cosmic expansion, the growth of structure and the linear bias to previous studies.

5.1 Configuration space measurements

5.1.1 Consistency between 3-multipole and 3-multipole

The correlation function multipoles and wedges of the eBOSS DR14 quasar sample with the weighting scheme $\mathcal{W}_{\rm focal-\mu}$ and the CLPT-GS model with parameters set to the best-fitting values are presented in figure 5.1. As mentioned previously, the error bars shown in the figure correspond to the diagonal elements of the inverse of the covariance matrix of the EZ mocks NGC+SGC using the same weighting scheme. The top panel represents the three Legendre multipoles used in this analysis and the botton panel the three wedges. The reference results for both analyses are displayed in table 5.1.

Legendre multipoles and wedges are just two decompositions of the same information, we thus expect both to be compatible and to provide similar uncertainty in the cosmological parameters. When analyzing the data under the same conditions (same RSD model, same fitting range, same covariance matrix), we confirm this behavior. The differences observed between the two methods are within one standard deviation, however the decomposition into Legendre multipoles provides the cosmological measurements with the best statistical precision. As a consequence, the constraints on $f\sigma_8$, H and D_A are obtained from the 3-multipole results. Indeed, their performance can be compared using the EZ mocks : we can check where the results of the fitting of the data sample stand compared to the dispersion of the results on the mocks. This is what figure 5.2 shows along with the measurements obtained from the data for the three redshift estimates ('z','$z_{\rm PCA}$','$z_{\rm MgII}$'), for the best-fitting values on the cosmological parameters (top panel) and for the 1σ uncertainty (bottom panel).

For the redshift estimate 'z', the results obtained from the data are similar w.r.t the distribution of the 1000 mocks. For the other redshift estimates, the results from the data deviate further from the spread of the mocks but one should bear in mind that the statistical uncertainty for these measurements can be much larger in the case of the 3-wedge analysis (see table 4.7). This behaviour is confirmed using the measurement of the uncertainties for the EZ mocks displayed in the bottom row of figure 5.2. The uncertainty in $\alpha_{\parallel}$ obtained with the 3-wedge approach when considering the redshift estimate 'z' is already shifted from the highest density region obtained from the mocks.

Figure 5.3 displays the likelihood contours with 68% and 95% confidence intervals of the reference results for a selection of pairs of parameters and for the two analyses using 3-multipole (orange) and 3-wedge (purple). We can see an important correlation between $\alpha_{\parallel}$ and $b\sigma_8$ which is consistent with the findings on the Outer Rim catalogs and on the

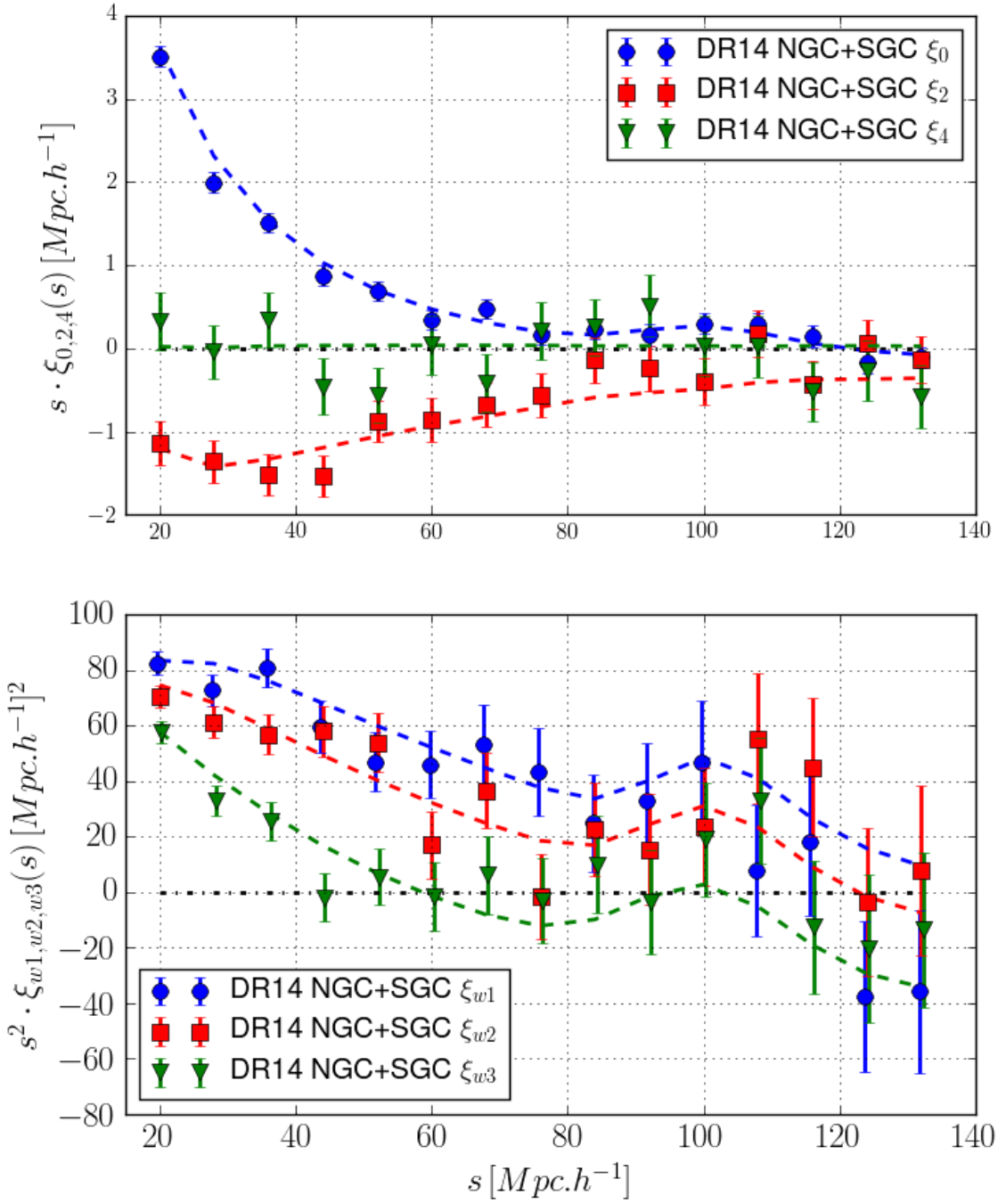


FIGURE 5.1 – Top panel : Monopole (blue) and quadrupole (red) and hexadecapole (green) of correlation function of the NGC+SGC eBOSS DR14 quasar sample fitted using the CLPT-GS model (dashed line) set to the best-fit parameters. Bottom panel : Same for the three wedges : $0< \mu <1/3$ (blue), $1/3< \mu <2/3$ (red) and $2/3< \mu <1$ (green).The fit is performed from 16 h^{-1}Mpc to 136 h^{-1}Mpc using binwidth of 8 h^{-1}Mpc. The covariance matrices are determined from the EZ mocks with a correction to equalize small differences in area between the footprint of the EZ mocks and the one of the data sample.

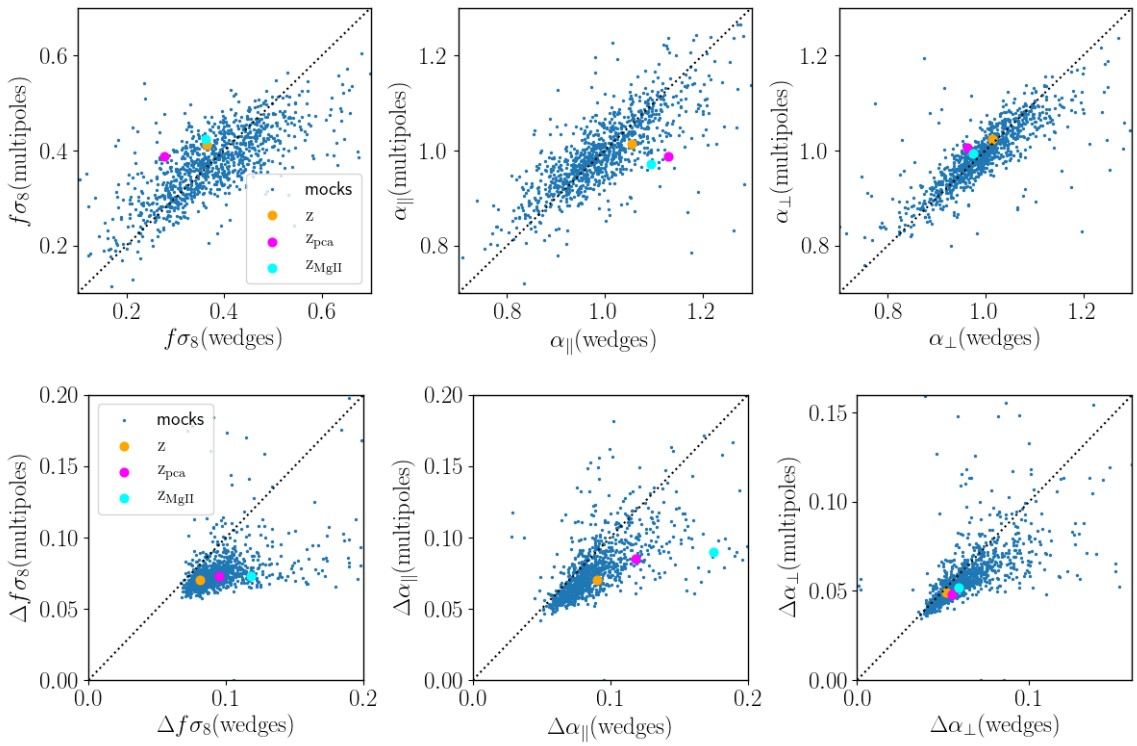


Figure 5.2 – Upper row : comparison between the 3-multipole and 3-wedge results on the cosmological parameters for the 3 redshift estimates and for the 1,000 EZ mocks. Bottom row : comparison of the uncertainties obtained for the two methods. Although one can note some differences between the 2 methods, they are consistently explained by the expected statistics.

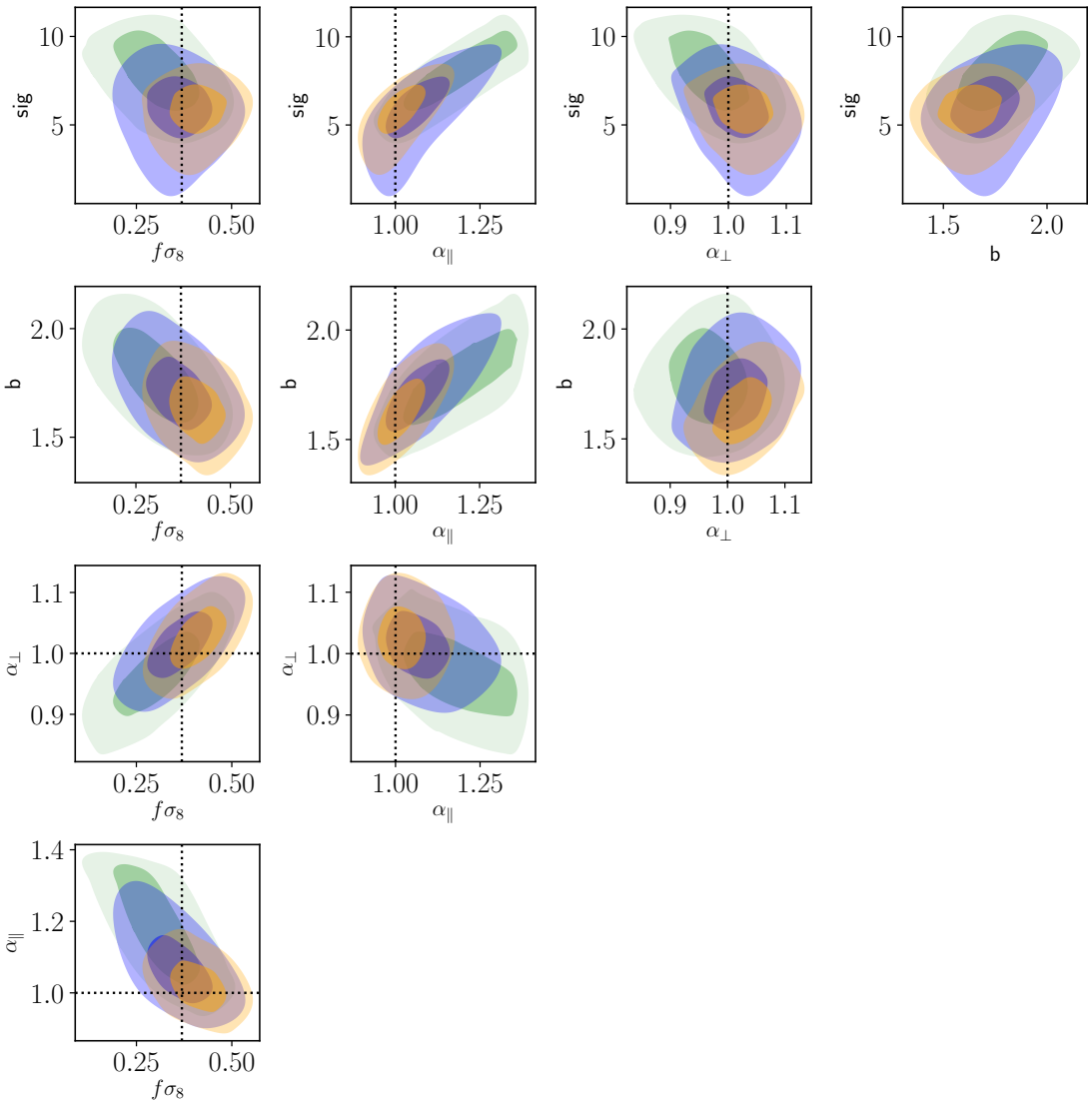


FIGURE 5.3 – Two-dimensional likelihood contours, showing the 68% (darker regions) and 95% (lighter regions) confidence intervals for various combinations of the parameters obtained from the anisotropic fit using 3-wedge (blue), 3-multipole (orange) and 2-multipole (green).

TABLE 5.2 – Results of the best fit parameters, and the statistical and systematical uncertainties for the 3-multipole analysis. The lower table shows the correlation coefficient between the 5 parameters in the RSD modeling.

parameter	best fit	statistical	systematic
$b\sigma_8$	1.038	$^{+0.060}_{-0.057}$	–
$f\sigma_8$	0.426	$^{+0.070}_{-0.070}$	0.033
$\alpha_{\rm par}$	1.012	$^{+0.071}_{-0.064}$	0.038
$\alpha_{\rm perp}$	1.031	$^{+0.050}_{-0.048}$	0.006
$\sigma_{\rm tot}$	5.94	$^{+1.19}_{-1.40}$	–

	$\alpha_\parallel$	$\alpha_\perp$	$b\sigma_8$	$f\sigma_8$	$\sigma_{\rm tot}$
$\alpha_\parallel$	1	-0.05	0.70	-0.38	0.68
$\alpha_\perp$		1	0.42	0.58	-0.15
$b\sigma_8$			1	-0.33	0.18
$f\sigma_8$				1	0.06
$\sigma_{\rm tot}$					1

data when performing tests on the bias prescription. We also see a significant correlation between $\alpha_\parallel$ and $\sigma_{\rm tot}$ which is consistent with the improvement of the precision on $\alpha_\parallel$ when fixing the redshift resolution in the fitting. The degeneracy between $f\sigma_8$ and the AP parameters demonstrates the importance of fitting them jointly in order to provide a measurement of the growth rate of structure independent of the fiducial cosmology. The green contours also show the results using the 2-multipole (without $\ell = 4$) to illustrate the sensitivity of the hexadecapole to the cosmological parameters. It highlights what we already found with the Outer Rim simulation : the hexadecapole is mostly sensitive to the variations of the geometrical parameters $\alpha_\parallel$ and $\alpha_\perp$. Adding this higher-order multipole in the fitting thus contributes to break the degeneracy between the growth rate and the AP parameters.

Finally, we conclude that the differences observed between the two methods are consistently explained by the expected statistics, and we consider the 3-multipole analysis as the results of this work. In Table 6.2, we summarize the results of this work and the correlation between the five parameters obtained from the 3-multipole analysis.

The measured dilation of scales using the eBOSS DR14 quasar sample, $\alpha_\parallel$ and $\alpha_\perp$, can be converted into cosmological parameters according to equations 4.10. We measured the expansion rate $H(z)$ and the angular diameter distance $D_A(z)$:

$$H(z_{\rm eff}) \cdot r_s(z_d) = 23.5^{+1.7}_{-1.9}\, 10^3\, {\rm km.s^{-1}} \quad (5.1)$$

$$D_A(z_{\rm eff})/r_s(z_d) = 12.58^{+0.61}_{-0.78} \quad (5.2)$$

where $r_s(z_d)$ is the comoving sound horizon at the end of the baryon drag epoch. In the case of the isotropic analysis, we measure : $\alpha_{\rm iso} = 1.021^{+0.046}_{-0.044}$. This isotropic shift of the BAO feature can be converted into the spherically averaged distance D_V using equation 4.11 :

$$D_V(z_{\rm eff} = 1.52)/r_s(z_d) = 26.8 \pm 1.1 \quad (5.3)$$

where all the quoted uncertainties include systematic and statistical contributions which are added in quadrature.

5.1.2 Consistency between single and multiple redshift bins

The results presented above have been obtained at the effective redshift of the quasar sample $z_{\rm eff} = 1.52$. However, this sample covers a wide redshift range probing different epochs of the universe. We thus expect a significant evolution with redshift of the fitting parameters, such as the bias and the redshift uncertainty, but also the growth rate. It means that the analyses that are performed at an effective redshift inevitably lead to loss of information while constraining the evolution of cosmological parameters with redshift will better constrain potential deviations from the standard cosmological model.

Therefore, the approach we adopt here is to split the sample into two distinct redshift bins with similar effective volume and redo the same analysis in each redshift bin :

name	redshift range	$z_{\rm eff}$	$V_{\rm eff}$ [Gpc^3]
lowz	$0.8 \leq z \leq 1.5$	1.19	0.126
highz	$1.5 \leq z \leq 2.2$	1.82	0.119

Figure 5.4 shows the three Legendre multipoles of the correlation function, monopole (left panel), quadrupole (middle panel) and hexadecapole (right panel), for the two redshift bins, *lowz* in blue and *highz* in red. The solid lines show the CLPT-GS model with parameters set to the best-fitting values.

For the monopole, the effect is mostly visible at small scales where the amplitude of the correlation function increases with redshift. Although the linear theory is no longer valid at these small scales, we remind the Kaiser formula for the monopole which is proportional to $(b\sigma_8)^2 + 2/3 b\sigma_8 f\sigma_8 + 1/5(f\sigma_8)^2$. It gives some intuition on the dependency of the monopole to $b\sigma_8$, which is a slightly increasing function with redshift, and $f\sigma_8$ which is a decreasing function with redshift (b and f may have a strong increasing dependence with redshift but σ_8 decreases with redshift which counterbalances their effect). The bias contribution is the dominant effect on the monopole so the global effect is an increase of the amplitude of the monopole.

The Kaiser formula for the quadrupole is proportional to $4/3 b\sigma_8 f\sigma_8 + 4/7(f\sigma_8)^2$ where the dominant dependency is with $f\sigma_8$ which decreases with redshift. We thus expect the overall absolute amplitude of the quadrupole to decrease with redshift as seen in the data and the best-fitting model. However, another significant effect which has already been highlighted is the impact of redshift uncertainty which is encoded by the fitting parameter σ_{tot} [1]. We expect the redshift uncertainties to increase with redshift as higher redshift quasars have spectra with lower signal-to-noise ration, making the determination of redshift noisier. We also showed in that it depends on the redshift estimate and comparison between different redshift estimates lead to the Gaussian modeling for the redshift uncertainty distribution given by equation 2.7. Therefore, the *highz* bin has a significant higher mean redshift uncertainty that causes the increase of the amplitude of the quadrupole at small scales.

For the hexadecapole, the Kaiser formula is proportional to $(f\sigma_8)^2$ with no bias contribution at large scales. Figure 4.8 also shows that the hexadecapole is mostly sensitive to the AP parameters. Both can have an impact on the redshift evolution, however given the statistical uncertainties that are large, we see no particular trend with redshift.

Table 5.3 presents the results of the fitting of the Legendre multipoles in the three redshift bins. We adopt the same prior on $\alpha_{\parallel,\perp}$ as in [3] in order to avoid secondary minima which could happen when fitting subsamples with smaller volumes. However, we

1. It also includes the intra-halo velocity dispersion related to the FoG effect but we showed in section 4.2.3 that it is sub-dominant in our fitting range

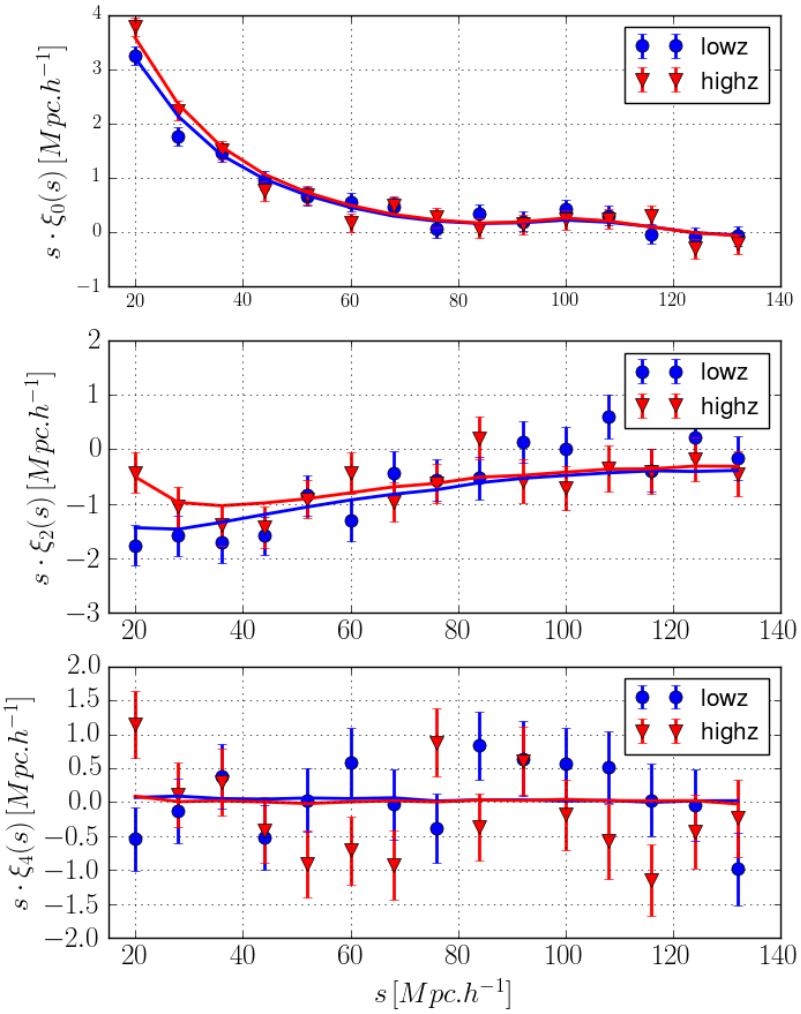


FIGURE 5.4 – Correlation function monopole (left panel), quadrupole (middle panel) and hexadecapole (right panel) for the two redshift bins of the eBOSS DR14 quasar sample : *lowz* in blue and *highz* in red. The solid lines represent the CLPT-GS model with parameters set to the best fitting values and the error bars are the diagonal elements of the inverse of the covariance matrix obtained from the 1,000 EZ mocks NGC+SGC.

TABLE 5.3 – Results for the anisotropic full-shape analysis

type	config	$b\sigma_8$	$f\sigma_8$	$\alpha_{\rm par}$	$\alpha_{\rm perp}$	$\sigma_{\rm totorP}$	χ^2/dof
CF	*lowz*	$0.893^{+0.073}_{-0.067}$	$0.398^{+0.093}_{-0.088}$	$1.041^{+0.082}_{-0.070}$	$1.048^{+0.091}_{-0.078}$	$4.24^{+1.82}_{-3.44}$	$40/(45-5)$
		$0.894^{+0.072}_{-0.068}$	$0.398^{+0.092}_{-0.087}$	$1.042^{+0.062}_{-0.056}$	$1.048^{+0.084}_{-0.074}$	4.24 (fixed)	$40/(45-5)$
	highz	$1.238^{+0.118}_{-0.109}$	$0.490^{+0.121}_{-0.124}$	$1.091^{+0.139}_{-0.120}$	$1.045^{+0.069}_{-0.060}$	$9.50^{+1.79}_{-1.63}$	$47/(45-5)$
		$1.237^{+0.100}_{-0.091}$	$0.489^{+0.120}_{-0.120}$	$1.091^{+0.077}_{-0.066}$	$1.044^{+0.070}_{-0.059}$	9.50 (fixed)	$47/(45-5)$
PS	*lowz*	0.900 ± 0.056	0.440 ± 0.087	0.994 ± 0.075	1.027 ± 0.075	4.11 ± 0.58	$71/(84-7)$
	highz	0.947 ± 0.077	0.468 ± 0.091	0.980 ± 0.130	1.039 ± 0.067	6.38 ± 0.77	$99/(84-7)$

TABLE 5.4 – Cosmological measurements obtained from the analysis of the two redshift bins and from the full redshift range.

config	$z_{\rm eff}$	$f\sigma_8$	$H(z_{\rm eff}) \cdot r_s(z_d) \times 10^3$	$D_A(z_{\rm eff})/r_s(z_d)$
lowz	1.19	$0.398^{+0.093}_{-0.088}$	$19.1^{+1.4}_{-1.4}$	$12.46^{+1.12}_{-1.10}$
highz	1.82	$0.490^{+0.121}_{-0.124}$	$25.3^{+3.1}_{-2.9}$	$12.64^{+1.19}_{-0.80}$
full	1.52	0.426 ± 0.077	$23.5^{+1.7}_{-1.9}$	$12.58^{+0.61}_{-0.78}$

also checked that when fitting with the same priors as the one given in table 4.1.1, we get the same results. Our results are compatible with those from the power spectrum analysis of the same sample [3] within one standard deviation.

config	parameter	prior
lowz	$\alpha_\parallel$	flat prior, range (0.5,1.5)
	$\alpha_\perp$	flat prior, range (0.5,1.5)
highz	$\alpha_\parallel$	flat prior, range (0.7,1.7)
	$\alpha_\perp$	flat prior, range (0.8,1.2)

As expected, σ_{tot} shows indicates a redshift evolution across the redshift range but in general for all parameters, the statistical uncertainty is too large to claim a significant redshift dependence. Moreover, as mentioned before, for $f\sigma_8$ and $b\sigma_8$ there is a counterbalance between the behavior of f (or b) with redshift and the one of σ_8. [3] also investigated the redshift evolution of the parameters across three redshift bins and also reported no significant redshift-dependence on the cosmological parameters given the current statistical precision. When fixing the redshift resolution to the best-fit value, we found that it improves the statistical precision on $\alpha_\parallel$ by almost a factor 2 for the *highz* bin. It supports the conclusion we found when fitting the full redshift range that understanding the spectroscopic redshift resolution of the quasar sample is necessary to reach the expected precision on the cosmological parameters.

We derive the cosmological constraints on $H(z_{\rm eff}) \cdot r_s(z_d)$ and $D_A(z_{\rm eff})/r_s(z_d)$ from the clustering measurements presented in table 5.3. The cosmological parameters obtained from the analysis of the two redshift bins and from the full redshift range are summarized in table 5.4. Although splitting the data sample into redshift bins allows us to measure the redshift evolution of the fitting parameters, the procedure is computationnally expensive and most of all, we loose some statistical precision compared to the full redshift bin.

An alternative method that avoids binning has been developped in [4] for BAO parameters and more recently adapted for RSD in [5]. Both validated their methods using mocks [6, 7]. They use a parametrization for the redshift evolution of the cosmological parameters using redshift weights, more details on the technique are provided in section 5.2.1.2.

Table 5.5 – Clustering analyses of the eBOSS DR14 quasar sample

reference	observables	redshift range	fitting range
Full-shape analysis			
[1]	$\xi_{l=0,2,4}$	$z_{\rm eff} = 1.52$	16-138 h^{-1}Mpc
[8]	$\xi_{l=0,2,4}$	$z_{\rm eff} = 1.52$	16-160 h^{-1}Mpc
[3]	$P_{l=0,2,4}$	$z_{\rm eff} = 1.52$	0.02-0.3 hMpc^{-1}
[9]	$P_{l=0,2}$	redshift-weight	0.001-0.3 hMpc^{-1}
[10]	$P_{l=0,2}$	redshift-weight	0.02-0.3 hMpc^{-1}
BAO-only analysis			
eBOSS BAO	ξ_0+P_0	$z_{\rm eff} = 1.52$	–
[11]	P_0	redshift-weight	–
[12]	P_0	redshift-weight	–

5.2 Consistency with other measurements with the DR14 sample

Our study complements the measurement of the BAO feature presented in Ata *et al.* [2] and is accompanied by five companion papers for full-shape analysis and two papers on BAO-only analysis. All studies are using the same DR14 quasar sample that is analysed under the same fiducial cosmology. Table 5.5 summarizes each reference, the observables in redshift space, the redshift range (effective redshift or a redshift-weighting technique to probe the redshift evolution of the cosmological parameters across the redshift range) and the fitting range.

First, we will present the consistency between the standard RSD analyses at the effective redshift of the sample $z_{\rm eff} = 1.52$ [1, 8, 3], then we will explain the redshift-weighting technique and how it compares with the approach using multiple redshift bins. Finally we will show the consensus among all the full-shape analyses of the eBOSS DR14 quasar sample.
Regarding the two additional BAO analyses, presented in [12, 11] and which complement the measurement of the spherically-averaged distance presented in [2], we will not present the consistency between the two methods but their comparison to [2] can be found in each paper.

5.2.1 Consistency between full-shape analyses

5.2.1.1 Comparison between standard analyses

The clustering analysis presented in this paper is based on the eBOSS DR14 quasar sample in the redshift range $0.8 \leq z \leq 2.2$, using Legendre multipoles with $\ell = 0, 2, 4$ and three wedges of the correlation function on the s-range from 16 h^{-1}Mpc to 138 h^{-1}Mpc. We use the CLPT predictions to take into account the non-linear evolution of the density and velocity fields with a Gaussian streaming model to perform the real-to-redshift space mapping. We find consistent results between the two methods although in our case the Legendre multipoles basis decomposition provides the cosmological measurements with the best statistical precision. So we use the constraints on the cosmological parameters obtained using the 3-multipole fit as our reference results which yield constraints on the

cosmological parameters $f\sigma_8(z_{\rm eff})$, $H(z_{\rm eff})$, and $D_A(z_{\rm eff})$ at the effective redshift $z_{\rm eff} = 1.52$.

Our companion paper in Fourier space [3] uses the power spectrum monopole, quadrupole and hexadecapole measurements on the k-range, $0.02 \leq k\,[h{\rm Mpc}^{-1}] \leq 0.30$, shifting the centres of k-bins by fractions of 1/4 of the bin size and averaging the four derived likelihoods. Their RSD modeling is based on the TNS model described in section 3.4.2 and the non-linear power spectra $P_{\delta\delta}$, $P_{\delta\theta}$ and $P_{\theta\theta}$ are computed using the 2-loop resumed perturbation theory [13]. Their bias modeling includes 2 Eulerian bias parameters (linear and non-linear) and 2 non-local bias parameters that can be fixed when assuming a local bias in Lagrangian space.

[8] perform an analysis using Legendre polynomials with order $\ell = 0, 2, 4$ and clustering wedges in configuration space as the work presented in this thesis. However, the fitting range extends to 160 h^{-1}Mpc and the RSD modeling is based on the power spectrum which they Fourier transform to obtain predictions for the two-point correlation function. They use a dispersion model extended to one-loop contribution developed by [14] and [15] along with a non-Gaussian FoG term such that :

$$P(k,\mu) = P_{\rm nonlin}(k,\mu)\, F_{\rm FoG}(k,\mu)\, e^{-(k\mu\sigma_{\rm zerr})^2} \tag{5.4}$$

where $P_{\rm nonlin}(k,\mu)$ is a non-linear extension of the Kaiser formula where $P_{\delta\delta}$, $P_{\delta\theta}$ and $P_{\theta\theta}$ are computed using the 2-loop renormalized perturbation theory including Galilean invariances of the equations of motion (gRPT, Crocce, Blas and Scoccimarro in prep) and $P_{\rm nonlin}(k,\mu)$ also includes contribution from the tree-level bispectrum (3-point statistics) and a quadratic linear theory power spectrum term. Their bias is modelled as described in [16], which includes both local and non-local contribution as the model used in [3]. As in [14], they model the effect of random motions on small-scales that give rise to the FoG effect by considering an exponential distribution :

$$F_{\rm FoG}(k,\mu) = \frac{1}{\sqrt{1 + f^2\mu^2k^2a_{\rm vir}^2}}\, \exp\left(\frac{-f^2\mu^2k^2\sigma_v^2}{1 + f^2\mu^2k^2a_{\rm vir}^2}\right) \tag{5.5}$$

where $a_{\rm vir}$ is treated as a free parameter and represents the kurtosis of the velocity distribution at small scales while the dispersion velocity σ_v is obtained from linear theory prediciton and treated as a constant with scale. Finally, the exponential term accounts for the redshift uncertainty in the quasar sample with a free parameter, $\sigma_{\rm zerr}$.

All the companion analyses use the weighting scheme based on $\mathcal{W}_{\rm focal}$ with a weight according to the inverse of the spectroscopic efficiency. In our work, we also discard the paircounts in the region $\mu > (1 - 1/480)$ to account for the effect of upweighting due to close pairs ($\mathcal{W}_{{\rm focal}-\mu}$). [3] report an observational systematic estimated from the EZ mocks using different weighting schemes that they add in quadrature to the statistical and modeling uncertainties. Figure 5.5 presents a comparison of the two-dimensional and one-dimensional posterior distribution of $b_1\sigma_8$, $f\sigma_8$, D_A/r_s and Hr_s at $z_{\rm eff} = 1.52$ for the two RSD analyses described above : in green for [8], in orange for [3] and this work in purple. The precision which is shown in the figure does not include the systematic error budget, each analysis reports a systematic uncertainty of up to ∼40% the statistical precision. Despite the differences in the modeling and fitting range, the three studies yield consistent results for the cosmological parameters. However the $b_1\sigma_8$ panels show a $\sim 1\sigma$ discrepancy between the analyses in Fourier space and configuration space. In our work, we investigated the effect of changing the prescription of the second-order bias F'' when evaluating the performance of the RSD model and we showed in section 4.2.3 that letting

this parameter vary could shift the first-order bias to lower values up to $\Delta b\sigma_8 = 0.038$. It would therefore reduce the difference we see with the Fourier space measurement but further investigations would be needed to test the effect of different bias models before providing a measurement of the bias for quasar clustering that could be used to constrain the astrophysical properties of quasar formation. We also checked that the difference in bias prescriptions did not affect the cosmological parameters of interest.

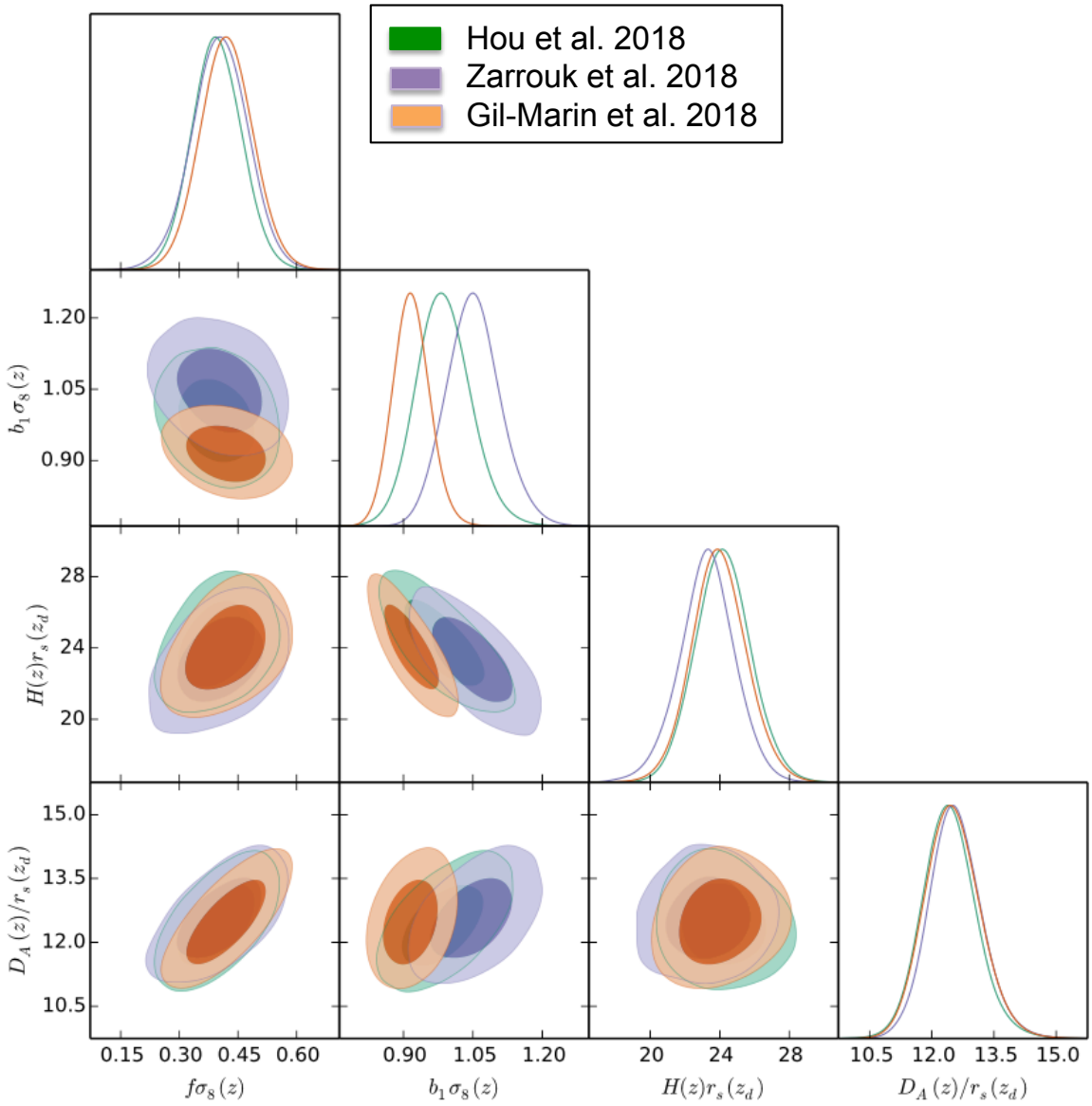


FIGURE 5.5 – Marginalized posterior distributions of the parameters $b_1\sigma_8$, $f\sigma_8$, D_A/r_s and Hr_s obtained at the effective redshift of the quasar sample $z_{\rm eff} = 1.52$ from measurements from this work [1] in purple, from [8] in green and from [3] in orange. The contours represent the 68% (darker regions) and 95% (lighter regions) confidence levels. From [3].

The two-dimensional posterior distributions can also be shown for the parameters $f\sigma_8$, D_v/r_s (equation 4.11) and F_{AP} (equation 4.12) along with the constraints inferred from the *Planck* CMB measurements [17] under the assumption of a flat ΛCDM cosmology. The CMB constraints on these parameters are in agreement with the results obtained from the clustering analyses of the eBOSS quasar sample, which shows the consistency of the DR14

measurements within the context of the ΛCDM model.

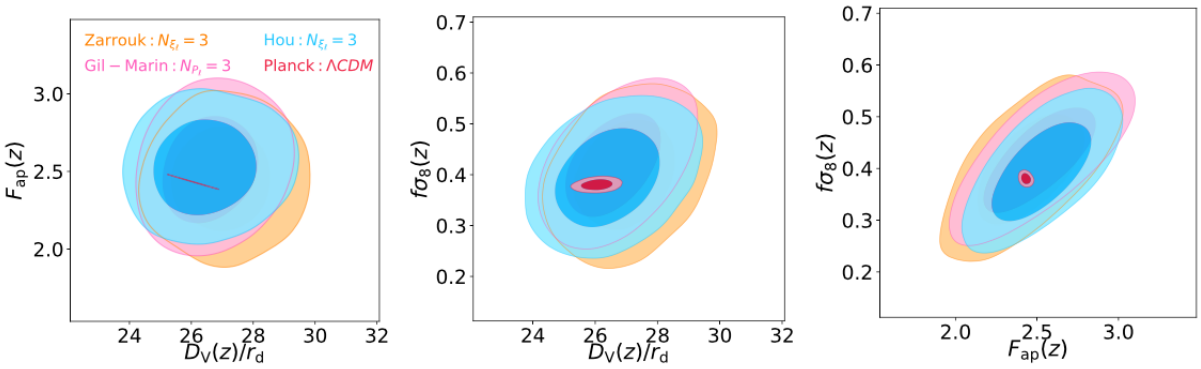


FIGURE 5.6 – Marginalized two-dimensional posterior distributions of the parameters $f\sigma_8$, D_v/r_s and F_{AP} obtained at the effective redshift of the quasar sample $z_{\rm eff} = 1.52$ from measurements from this work [1] in orange, from [8] in blue and from [3] in pink. The contours represent the 68% (darker regions) and 95% (lighter regions) confidence levels. The red contour shows the *Planck* prediction for a flat ΛCDM cosmology. From [8].

5.2.1.2 Redshift-weighting technique

Standard analyses split the redshift range into different redshift bins and repeat the traditional analysis inside each bin in order to take into account the redshift evolution of the cosmological parameters across the redshift range. However, this method is computationally expensive and causes loss of information from galaxy pairs into different bins. Moreover, it decreases the signal-to-noise ratio and low signal-to-noise detections can lead to non-Gaussian likehood profiles for the cosmological parameters. Instead of splitting the sample into multiple redshift slices, [4] proposed to compress the information contained in the redshift direction onto a small number of 'modes' from which they derived a set of redshift weights for BAO measurements and applied the procedure on mock catalogs [6]. Then, [5] extended the technique for RSD measurements and also tested it on mock catalogs [7] while [18] proposed optimal redshift weights for non-local bias at large scales that can probe primordial non-Gaussianities.

All these methods are based on a linear data compression as derived originally in [19] where the uncompressed data vector in n z-bins can be written as :

$$\mathbf{X}(r) = \begin{pmatrix} x(r, z_1) \\ x(r, z_2) \\ \vdots \\ x(r, z_n) \end{pmatrix}$$

assuming that the n-dimensional data set $\mathbf{X}$ is Gaussian distributed with mean μ and covariance C. In the papers mentioned above, x refers to the two-point statistics ξ or $P(k)$. By weighting the redshift slices using the n-dimensional vector $\mathbf{w} = (w_1, w_2, \ldots, w_n)$, the data vector is linearly compressed into one single z-bin according to these weights so that :

$$x_w(x) = \mathbf{w}^T \mathbf{X}(r) \tag{5.6}$$

where x_w is the weighted data set with mean $\mathbf{w}\mu$ and variance $\mathbf{w}^T C\mathbf{w}$. Then by definition, the optimal weights are the ones that minimize the uncertainty in the parameters of interest θ_i. These uncertainties can be obtained using the elements of the Fisher matrix. We remind that the Fisher information matrix F is defined as the second derivative of the logarithmic likelihood function L w.r.t. the set of parameters to be measured :

$$F_{i,j} = \langle \frac{\partial^2 L}{\partial\theta_i \partial\theta_i} \rangle \tag{5.7}$$

In Fourier space the covariance matrix C of the power spectrum in the absence of a survey window is diagonal (for the analysis of the eBOSS DR14 sample, [3] shows that it remains mostly diagonal and the off-diagonal terms account for few %). Then, for each redshift slice of volume dV and expected target density $\bar{n}(r)$, the covariance matrix is :

$$C \sim (P + \frac{1}{\bar{n}})^2 \frac{1}{dV} \tag{5.8}$$

Under these assumptions, one can derive a set of optimal weights that maximize F_{ii} w.r.t to $\mathbf{w}$.

The procedure to include redshift-weights can be summarized as follows :

(i) Perform a Taylor series around the cosmological parameter of interest measured at the effective redshift of the sample as a function of optimal weights

(ii) Linear compression of the data set $y = w^T\mathbf{x}$

(iii) Search for optimal weights by minimising the uncertainty in this parameter using the Fisher matrix

(iv) Reconstruct the redshift-evolution of each set of parameters

(v) Apply weights to the data and include them in the model

To study the evolution of the distance-redshift relation with redshift, [4] expanded the comoving distance $\chi(z)$ using a Taylor series around the fiducial value at the effective redshift (also called pivot redshift) of the sample :

$$\chi(z) = \chi_{fid}(z)\, \alpha_0[1 + \alpha_1 x + \frac{1}{2}\alpha_2 x^2] \tag{5.9}$$

where $1+x = \chi_{fid}(z)/\chi_{fid}(z_{\rm eff})$ and the parameters $\alpha_{0,1,2}$ correspond to the set of optimal weights to be found by maximising F w.r.t χ.
In [5], they proposed to expand $\Omega_m(z)$ around the fiducial model and this parametrization has been used for the analysis of the eBOSS DR14 sample as the interest is to explore deviations from ΛCDM-model both in terms of geometrical and gravitational effects. In [9], they also include a parametrization of the growth rate alone as it represents a more direct way to measure deviations from the underlying framework for gravity, and the geometrical evolution is thus kept separate. Both parametrizations as Taylor expansions yield :

$$\Omega_m(z) = \Omega_{m,fid}(z)\, q_0[1 + q_1 y(z)] \tag{5.10}$$

$$(f\sigma_8)(z) = (f\sigma_8)_{fid}(z)\, p_0[1 + p_1 g(z)] \tag{5.11}$$

where $y(z) = \Omega_{m,fid}(z)/\Omega_{m,fid}(z_{\rm eff})$ and $g(z) = (f\sigma_8)_{fid}/(f\sigma_8)_{fid}(z_{\rm eff})$.

[10] propose an alternative approach to extract the information in redshift which is also based on a power spectrum analysis using the monopole and the quadrupole only (in the k-range of $0.02 \leq k\,[h\mathrm{Mpc}^{-1}] \leq 0.30$). They construct an optimally redshift-weighted sample and compare to a power spectrum template based on the regularised perturbation theory

up to second order. Using four redshift-weighted power spectra, they constrain $\alpha_\perp, \alpha_\parallel$ and $f\sigma_8$ at four effective redshifts (0.98, 1.23, 1.53 and 1.94). Contrary to the analysis presented in [9] where the whole redshift range is considered using weighted multipoles, [10] derive redshift weights that divide the sample into optimal z-bins and perform the same analysis in each z-bin. It is thus a hybrid between the redshift-weighting technique that tracks the redshift evolution across the whole redshift range and standard analyses.

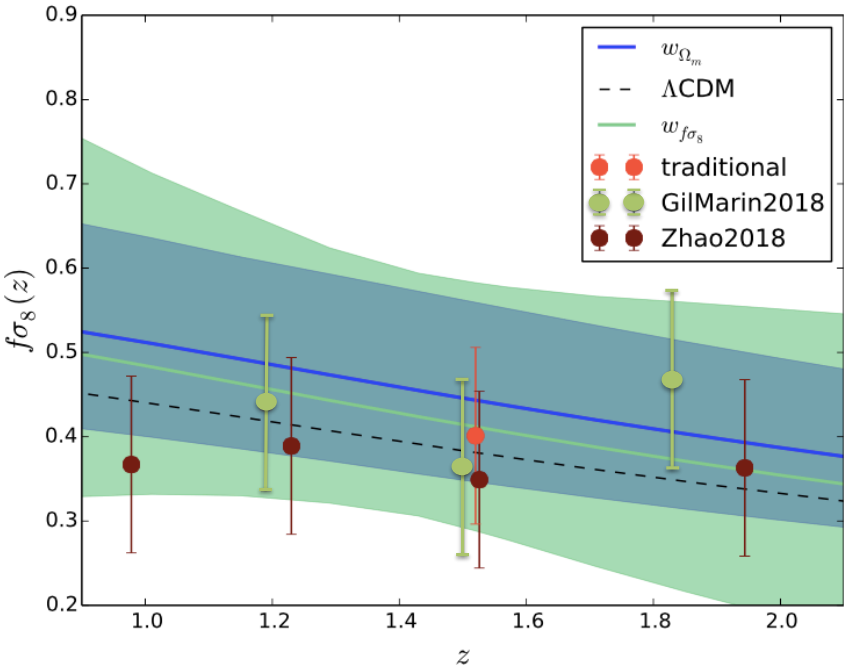


FIGURE 5.7 – Evolution of $f\sigma_8$ as a function of redshift obtained by [9] using redshift-weights derived for Ω_m in blue and for $f\sigma_8$ in green, by [10] in dark red and by [3] in magenta after splitting the sample into three redshift bins. For comparison, the measurement at the effective redshift of the full sample $z_{\rm eff} = 1.52$ is shown in orange. From [9].

Figure 5.7 shows the comparison of the $f\sigma_8$ evolution obtained by different analyses : the optimal redshift-weights derived for Ω_m in blue and the optimal redshift-weights derived for $f\sigma_8$ in green where both come from [9], the four measurements in dark red from the redshift-weighting technique in [10] and the Fourier space measurements [3] in magenta by splitting the data in three redshift bins while the orange point is the measurement at the effective redshift $z_{\rm eff} = 1.52$. All methods agree and provide comparable uncertainties at this stage of the analysis. However, the redshift range of the quasar sample is particularly wide so we expect the redshift-evolution of the parameters to be more important. The quasar density is also low compared to the BOSS LRG sample, so splitting the data into redshift bins can have a significant impact of the signal-to-noise ratio, which is avoided when using the redshift-weights. For all these reasons, we expect this optimal redshift weighting technique to provide tighter constraints on the final eBOSS sample compared to the standard approach and the technique will be even more useful for the upcoming generation of surveys that will extend further the dynamical redshift range and reduce the statistical error by an order of magnitude at least.

5.2.1.3 Consensus between the RSD analyses

In order to do a consistent comparison between analyses at effective redshift and with redshift weights, the comparison presented in this section uses the results from the traditional analysis performed at an effective redshift for all studies. For redshift-weigthing techniques, it thus means that only FKP weights are taken into account as they correspond to the limit when there is no redshift dependence of the cosmological parameters. Moreover, for these 2 analyses the results come from the fitting of the first two even multipoles of the power spectrum. Finally, the comparison presented in this section uses the traditional weighting scheme, $\mathcal{W}_{\rm focal}$ for the four companion paper, except our work that also takes into account the effect of up-weighting the nearest neighbor quasar when there is a collision $\mathcal{W}_{\rm focal-\mu}$.

The likelihood contour constraints for the cosmological parameters $f\sigma_8$, $H(z)r_s$, and $D_A(z)/r_s$ at $z_{\rm eff} = 1.52$ for the five analyses described above are shown in figure 5.8. Each analysis uses a different model for the 2-point statistics, three are in Fourier space and two in configuration space. Despite those differences, there is good agreement between all analyses. These contours only show the statistical precision which is also similar. The one-dimensional likelihood for each parameter better displays the consistency between the measurements. For the three traditional analyses [1, 3, 8] the agreement is excellent. The systematic errors, which are not included in these contours, are estimated by the different groups and found to be up to 40% of the statistical precision.

The likelihood distribution for the two different redshift-weighting techniques [9, 10] when using no redshift-dependent weights are slightly wider but remain consistent with the others. In fact, the results from the analyses using redshift weights are obtained by fitting the monopole and quadrupole only. Adding the hexadecapole provides additional information that increases the sensitivity of the clustering observables to the cosmological parameters. We report no results using the first two even multipoles but we found that adding the hexadecapole could improve the statistical precision by few percents which is consistent to what is reported on table 9 of [3] in Fourier space. We refer the reader to Section 5 of each paper for additional information on the different approaches and on the comparison between the redshift-dependent weights and the traditional analysis at a singe effective redshift on the data.

5.2.2 Consistency with BAO-only analysis

All the analyses described in the previous section aim at measuring simultaneously the AP parameters ($\alpha_\parallel$, $\alpha_\perp$) and the rate of structure growth by modeling the anisotropic 2-point statistics. However, the position of the BAO peak can also be measured using the monopole only without considering either the full-shape of the monopole nor the higher multipoles that contain more non-linearities that are challenging to model. The BAO signature has now been detected in many different galaxy surveys probing the epoch $z < 1$, but also in Ly-α samples (at $z \sim 2.5$) using a variety of methods to analyze the density field. The analysis of the eBOSS DR14 quasar sample provides the first detection of the BAO peak in the intermediate redshift range $1 < z < 2.2$. The BAO-only analysis is presented in [2] using the same procedure as in [20] for BOSS DR10 and DR11 LRG sample and in [21] for BOSS DR9 LRG sample. The procedure is the same for both the correlation function and the power spectrum, except for the clustering estimator which is more complicated in Fourier space. In what follows, we summarize the main steps using the same notation as in [2] :

(i) computation of the two-point statistics (using the LS estimator for the correlation

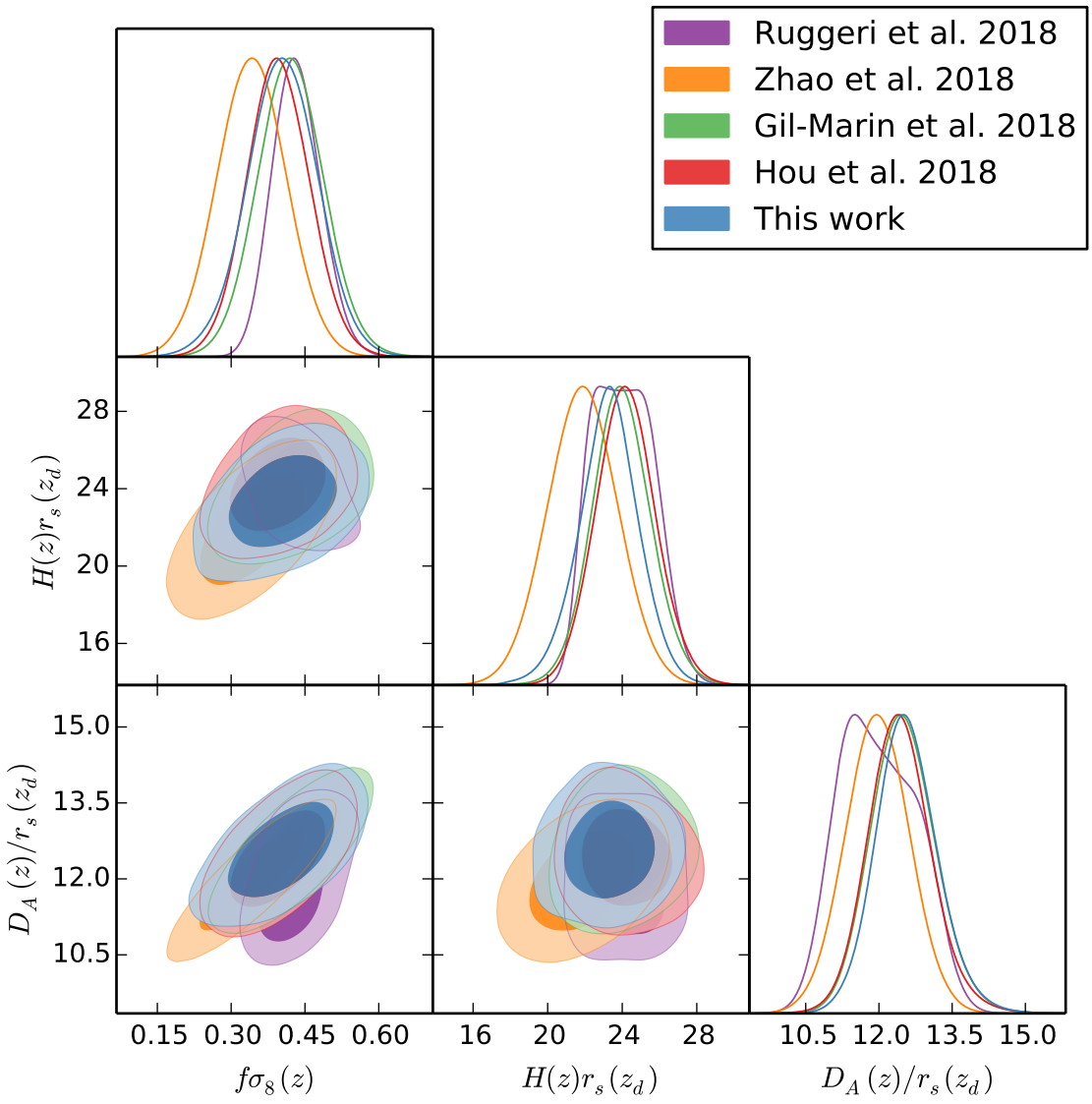


FIGURE 5.8 – Parameter contours for $f\sigma_8$, D_A and H for the predictions by the 5 companion papers using the same DR14Q dataset for traditional RSD analyses. Blue contours show the results presented in this work in configuration space, and red contours show the predictions by [8] in configuration space too using a second RSD modeling. The Fourier Space based analyses are shown in green contours for the results by [3] using a third RSD modeling, in magenta contours for the results by [9] and in orange contours for [10], both using redshift weighting techniques but with a different model.

function as described in section 2.7.2)

(ii) generate a template BAO feature $\xi_{\rm temp}$ using the linear power spectrum $P_{\rm lin}(k)$ obtained from CAMB assuming a fiducial cosmology

(iii) generate a template without the BAO feature $P_{\rm nw}(k)$ (where no BAO signal corresponds to "no wiggle" in k-space) obtained from the fitting formulae in [22] using the same fiducial cosmology

We can then model the correlation function following [23] :

$$\xi^{mod}(s) = B_0 \xi_{\rm temp}(\alpha, s) + A_1 + A_2/s + A_3/s^2 \tag{5.12}$$

where B_0 is a multiplicative constant allowing for an unknown large-scale bias and $A_{1,2,3}$ are the coefficients of the additive polynomial function to make the results insensitive to shifts in the broad-band shape of the measured monopole. The BAO template for the correlation function is obtained by Fourier transformation of the power spectrum :

$$\xi_{\rm temp}(s) = \int \frac{k^2 dk}{2\pi^2} P_{\rm temp}(k) j_0(ks) e^{-k^2 a^2} \tag{5.13}$$

where the Gaussian term has been introduced to damp oscillatory patterns associated with the Bessel function j_0 at high-k and induce better numerical convergence [20]. The exact damping scale is not important, in this analysis it is set to $a = 1h^{-1}$Mpc. The template for the power spectrum is given by :

$$P_{\rm temp}(k) = P_{\rm nw}(k) \left[1 + \left(\frac{P_{\rm lin}(k)}{P_{\rm nw}(k)} - 1 \right) e^{\frac{1}{2}k^2 \Sigma_{nl}^2} \right] \tag{5.14}$$

where the BAO signature in linear theory is described by the oscillatory pattern in the $O_{\rm lin}(k) = P_{\rm lin}(k)/P_{\rm nw}(k)$ and the Σ_{nl}^2 term is used to damp the acoustic oscillations in the linear theory power spectrum to account for the effects of non-linear evolution of the density field. In [2], we use $\Sigma_{nl}^2 = 6\,[h^{-1}\text{Mpc}]^2$ and they show that the results are insensitive of this choice, as also confirmed in galaxy samples.

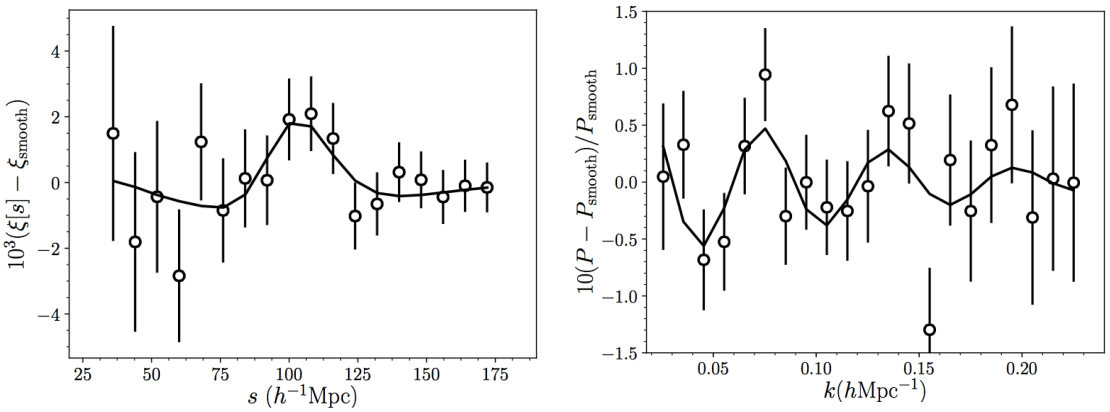


FIGURE 5.9 – Measurement of the spherically-averaged BAO signal in the monopole of the two-point correlation function (left) and power spectrum (right) using the eBOSS DR14 quasar sample. The smooth component of the best-fit model has been substracted from the best-fit model and data in order to isolate the BAO feature.

Figure 5.9 shows the measurement of the spherically-averaged BAO signal in the monopole of the two-point correlation function (left) and power spectrum (right) using the

eBOSS DR14 quasar sample. Each clustering static prefers the model with the BAO feature to the smooth model at better than 2.8σ according to the likelihood and detection significance shown in the right panel of figure 5.10. The dashed curve represents the model without BAO and the black curve represents the eBOSS quasar BAO distance measurement from the combination of correlation function and power spectrum, which is the final measurement of [2]. It yields a BAO measurement with a precision of 4.1%, which is shown in the right panel of figure 5.10, along with galaxy measurements at lower redshifts using the BOSS DR12 galaxy sample [24], the SDSS MGS, WiggleZ and 6dFGRS. At higher redshift, the constraint comes from the combination of the Ly-α forests auto-correlation [25] and the cross-correlation of Ly-α forests and quasars [26].

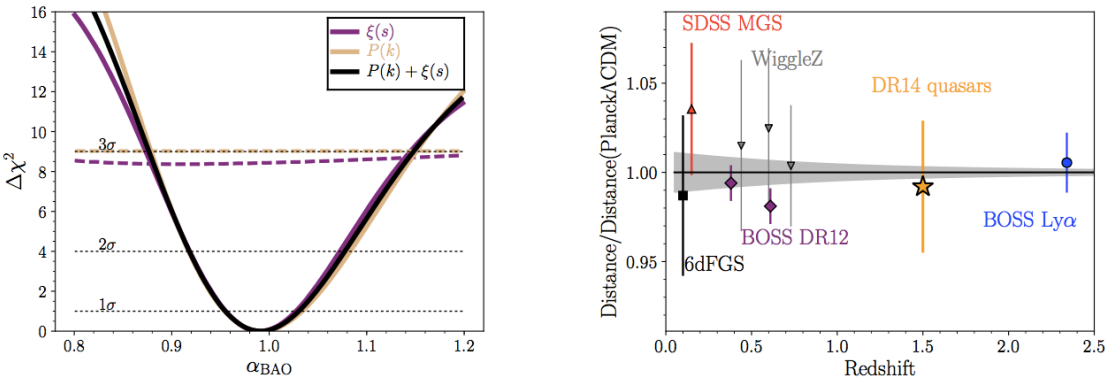


FIGURE 5.10 – Left panel : Likelihood of the isotropic BAO parameter $\alpha_{\rm iso}$ in terms of $\Delta\chi^2$ recovered from the fit of the data with a model that contains the BAO feature. The dashed line displays the same likelihood for the model without BAO where $\Delta\chi^2$ is determined by substracting the minimum χ^2 from the model with BAO. Right panel : Spherically-averaged BAO distance measurements compared to *Planck* ΛCDM prediction and extrapolated 68% confidence level. The eBOSS DR14 quasar sample measurement is represented by a gold star.

Table 5.6 summarizes the measurements of the isotropic shift of the BAO position from the monopole only (top of the table) and from the analysis of the three multipoles (bottom panel). The value and uncertainty of $\alpha_{\rm iso}$ from the '3-multipole full-AP' case is given by

$$\alpha_{\rm iso} = \alpha_{\parallel}^{1/3}\alpha_{\perp}^{2/3} \tag{5.15}$$

$$\sigma_{\alpha_{\rm iso}} = \sqrt{\Delta\alpha_{\parallel}\left(\frac{\partial\alpha_{\rm iso}}{\partial\alpha_{\parallel}}\right)^2 + \Delta\alpha_{\perp}\left(\frac{\partial\alpha_{\rm iso}}{\partial\alpha_{\perp}}\right)^2} \tag{5.16}$$

where $\Delta\alpha_{\parallel}$ and $\Delta\alpha_{\perp}$ are the uncertainty in $\alpha_{\parallel}$ and $\alpha_{\perp}$ respectively. All the quoted uncertainties include systematic and statistical contributions.

5.3 Combining measurements

In this section, we combine our measurements in configuration space with the measurements in Fourier space in [3]. Both are mathematical representation of the same underlying physics so we expect them to be highly correlated. The degree of correlation depends on how far these approaches are sensitive to different systematics.

Table 5.6 – Measurement of the spherically-averaged BAO distance D_V from BAO-only [2] and full-shape analysis [1] of the eBOSS DR14 quasar sample.

analysis	$\alpha_{\parallel}$	$\alpha_{\perp}$	$\alpha_{\rm iso}$	$D_V(z_{\rm eff})/r_s$
BAO-only $\xi_0 + P_0$	–	–	0.993 ± 0.038	26.0 ± 1.0
3-multipole full-AP	1.012 ± 0.068	1.031 ± 0.049	1.021 ± 0.042	26.8 ± 1.4
3-multipole isotropic	–	–	1.021 ± 0.045	26.8 ± 1.1

We follow the methodology given in [27] where for simplicity we assume a linear combination of the measurements of the individual probes such that the value of the combined measurement is given by

$$x^{\rm combined} = \sum_{i=1}^{N_{\rm probes}} w_i x_i \tag{5.17}$$

where the variable x stands for $f\sigma_8$, H or D_A and the weight w_i that is assigned to each measurement depends only on the covariance matrix among the x_i elements of the different probes. Here $N_{\rm probes} = 2$. The variance of the combined measurement is given by

$$\sigma^2_{x^{\rm combined}} = \sum_{ij} w_i w_j C_{ij} \tag{5.18}$$

where C_{ij} is the ij-element of the covariance matrix.

In order to get a combined measurement with the best uncertainty, the weights must satisfy a condition of minimum variance estimator. Provided that the weights are normalized (i.e. $\sum_i w_i = 1$), it means that by minimizing $\sigma^2_{x^{\rm combined}}$ w.r.t. w_i, we obtain

$$w_i = \frac{\sum_k (C^{-1})_{ik}}{\sum_{jk} (C^{-1})_{jk}} \tag{5.19}$$

The box below proposes an analytical derivation of the weights for two correlated parameters between two different probes that measure this parameter only. In practice, full-shape analyses measure simultaneously $f\sigma_8$, H and D_A and there are correlations between the three parameters within one method.

Combine two single measurements of the same dataset

We consider two measurements with their total uncertainty, $\chi_1 \pm \sigma_1$ and $\chi_2 \pm \sigma_2$, obtained from different methods applied to the same data sample. Both measurements are correlated with a factor of correlation ρ. We want to find the best way of combining both measurements in order to maximize the gain on the total uncertainty. Let x denote the combined value and σ_x its combined uncertainty :

$$x = \lambda\chi_1 + (1-\lambda)\chi_2 \tag{5.20}$$

$$\sigma_x^2 = \lambda^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2) - 2\lambda(\sigma_2^2 - \rho\sigma_1\sigma_2) + \sigma_2^2 \tag{5.21}$$

where λ is the coefficient to be found in order to optimize the combination, i.e. when σ_x^2 is minimal w.r.t. λ :

$$\frac{d\sigma_x^2}{d\lambda} = 0 \tag{5.22}$$

$$\Leftrightarrow 2\lambda(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2) - 2(\sigma_2^2 - \rho\sigma_1\sigma_2) = 0 \tag{5.23}$$

$$\Leftrightarrow \lambda = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \tag{5.24}$$

We note that when $\sigma_1 = \sigma_2$, $\lambda = 1/2$ as expected. Equation describes the way to optimize the combination of both measurements to obtain the best gain on the combined uncertainty. Then, by replacing equation 5.3 into equation 5.21, it gives for σ_x^2 :

$$\sigma_x^2 = \frac{\sigma_1^2\sigma_1^2 - (\rho\sigma_1\sigma_2)^2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \tag{5.25}$$

where the only unknown in this equation is ρ which is usually estimated from the covariance of the measurements in mock catalogs.

We apply this method to the eBOSS DR14 quasar sample. Figure 5.11 shows the correlation coefficients between the cosmological parameters $f\sigma_8$, $\alpha_\parallel$ and $\alpha_\perp$ from the two-point correlation function (CF) [1] and the power spectrum (PS) [3] for the EZ mocks with $0.8 < \alpha_{\parallel,\perp} < 1.2$. To compute the weights given by equation 5.19, we take the inverse of the covariance matrix for the ij-elements and we take the individual uncertainties σ_{ii} from the data such that $(C^{-1})_{ii} = \sigma_{ii}^2$.

The results on the cosmological parameters are summarized in table 5.7. At this stage of the analysis, combining Fourier space and configuration space measurement provides the following gain on the constraints put on the cosmological parameters : 6% on $\sigma_{f\sigma_8}$, 11% on σ_H and 7% on σ_{D_A}.

5.4 Comparison to previous works

5.4.1 Cosmological distances measurements

Table 6.3 summarizes the measurements of cosmological distances from BAO-only and full-shape analyses using different large-scale structure surveys. The results from BOSS

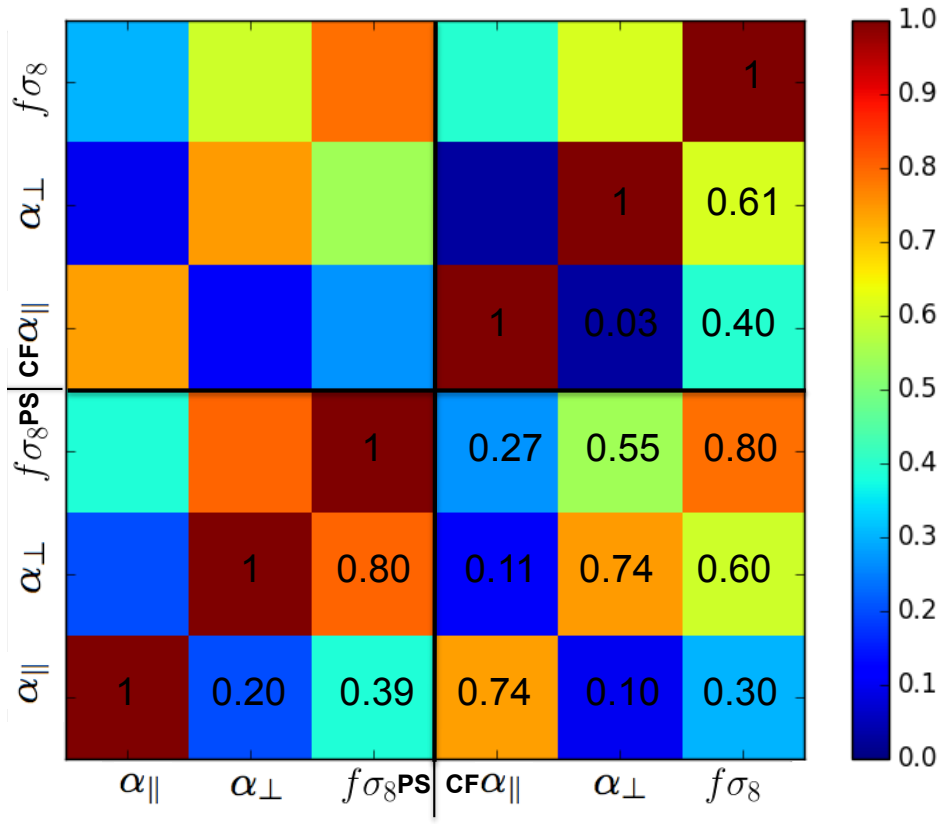


FIGURE 5.11 – Correlation coefficients between power spectrum (PS) and correlation function (CF) measurements of $f\sigma_8$, $\alpha_\parallel$ and $\alpha_\perp$ for the EZ mocks with $0.8 < \alpha_{\parallel,\perp} < 1.2$.

TABLE 5.7 – Measurement of the growth rate $f\sigma_8$, expansion rate H and angular distance D_A in configuration and Fourier space separately, along with the combination of both.

analysis	$f\sigma_8(z_{\rm eff})$	$H(z_{\rm eff}) \cdot r_s(z_d) \times 10^3$	$D_A(z_{\rm eff})/r_s(z_d)$
$\xi(s)$ [1]	0.426 ± 0.077	23.5 ± 1.8	12.58 ± 0.70
$P(k)$ [3]	0.420 ± 0.076	24.0 ± 1.8	12.48 ± 0.71
$\xi(s) + P(k)$	0.423 ± 0.071	23.8 ± 1.6	12.57 ± 0.65

TABLE 5.8 – Constraints on the BAO distances from BAO-only analyses using various large-scale structure surveys.

survey	reference	z	D_V	D_M	H
6dFGS BAO	[29]	0.106	4.5%	–	–
WiggleZ BAO	[30]	0.44	4.8%	–	–
		0.60	4.5%	–	–
		0.73	3.4%	–	–
WiggleZ FS	[31]	0.44	–	9.4%	9.4%
		0.60	–	6.9%	6.9%
		0.73	–	7.0%	7.2%
SDSS MGS BAO	[32]	0.15	3.8%	–	–
SDSS BOSS galaxies BAO+FS	[24]	0.38	1.09%	1.5%	2.4%
	[24]	0.51	1.01%	1.4%	2.2%
	[24]	0.61	1.03%	2.8%	2.2%
SDSS BOSS Ly-α BAO	[25]	2.40	–	5.8%	3.4%
	[33]	2.33	–	3.9%	2.8%
	combination	2.40	–	3.0%	2.0%
SDSS eBOSS LRG BAO	[28]	0.72	2.6%	–	–
SDSS eBOSS quasars BAO	[2]	1.52	3.8%	–	–
SDSS eBOSS quasars FS	[1]	1.52	4.1%	5.5%	7.9%

DR12 provided the best constraints from far with 1-3% precision on the three BAO distances at redshifts 0.38, 0.51, 0.61 using galaxies and at $z = 2.4$ using Ly-α forests. Our measurements from the monopole only and the full correlation function of the eBOSS quasar sample are shown at the end of the table. The statistics of our sample is an order of magnitude lower compared to BOSS DR12 galaxies but we stress that our measurement of the expansion rate D_A at $z = 1.52$ is competive with the one from the auto-correlation of Ly-α forests for instance. Compared to other surveys, our measurements are competitive with current constraints provided that in addition our measurements extend the constraints at the intermediate redshift range $1 < z < 2$. Another constraint on D_V from eBOSS arises using the eBOSS LRG sample which includes 80,118 LRGs. By combining them with the high-redshift tail of the BOSS galaxy sample at $z > 0.6$, [28] found a 2.6% measurement of the D_V at an effective redshift $z = 0.72$.

Figure 6.8 presents our measurements of the cosmological distances compared with the prediction of Λ-CDM using *Planck* results [17]. Also shown are the results of previous measurements : 6dFGS from [29], SDSS MGS from [32], BOSS DR12 from [24] and BOSS Lyα from the combination of the DR12 Lyα auto-correlation [25] and the cross-correlation of the Lyα forests [33]. The WiggleZ survey volume also overlaps BOSS and its measurements are less constraining than the ones of BOSS so we decide not to show them for clarity in the figure.

Our measurements are consistent with previous analyses and all measurements agree with the expansion history predicted by the Λ-CDM+GR concordance model using *Planck* measurements of the cosmological parameters.
We also compare the measurement of the spherically-averaged BAO distance between full-shape analysis (this work) and BAO-only [2]. The two measurements are in agreement and they provide similar constraints on this parameter (3.8% precision using BAO-only and 4.1% using full-shape correlation function).

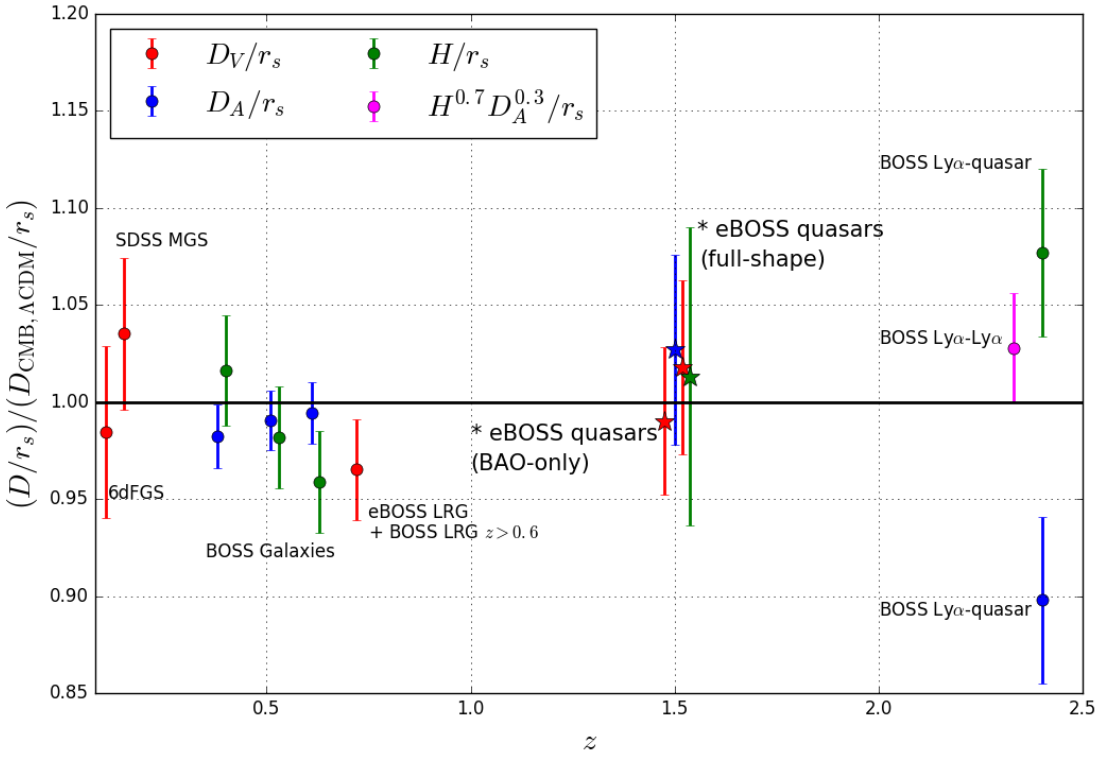


Figure 5.12 – Evolution of the BAO distances with redshift compared to the prediction from the flat Λ-CDM model with *Planck* parameters. The Hubble distance D_{H} is related to the Hubble parameter H by $D_{\mathrm{H}} = c/H$ and $D_{\mathrm{M}} = (1+z)D_{\mathrm{A}}$ where D_{M} is the comoving angular diameter distance. The BAO results from this work using the eBOSS DR14 quasars are represented by the * marker and are compared to previous analyses using galaxies and Ly-α forests to probe different epochs.

TABLE 5.9 – Constraints on $f\sigma_8$, D_A and H from full-shape analyses using various large-scale surveys. Note that when no values for D_A and H are quoted, the measurement of $f\sigma_8$ has been performed without marginalizing over the AP parameters.

survey	reference	z	$f\sigma_8$	AP
2dFGRS	[35]	0.17	13%	no
6dFGS	[36]	0.067	13%	no
WiggleZ	[31]	0.44	19.4%	yes
		0.60	16.2%	yes
		0.73	16.5%	yes
VVDS	[34]	0.77	36.7%	no
VIPERS	[37]	0.60	21.8%	no
		0.86	27.5%	no
Subaru FastSound	[38]	1.4	25%	no
SDSS MGS	[39]	0.15	30.6%	yes
SDSS BOSS galaxies FS	[24]	0.38	9.5%	yes
		0.51	8.7%	yes
		0.61	8.8%	yes
SDSS BOSS galaxies BAO+FS	[24]	0.38	9.2%	yes
	[24]	0.51	8.3%	yes
	[24]	0.61	8.0%	yes
SDSS eBOSS quasars	[1]	1.52	18.1%	yes

5.4.2 Growth rate measurements

The measurement of the anisotropic clustering of the DR14 eBOSS quasar sample produces the constraint on $f\sigma_8(z_{eff} = 1.52) = 0.426 \pm 0.079$ that is presented in Figure 6.9. The result is obtained from a fit of the $l = 0, 2, 4$ Legendre multipoles of the correlation function, and the uncertainty includes systematic errors due to the modeling of the RSD and statistical contributions added in quadrature. As originally highlighted in [34], the measurement of the growth rate of structure can be a direct test of GR. Our results confirm the validity of GR in the intermediate redshift range ($1 < z < 2$) probed by eBOSS quasars and there is consistency between our result and the measurement done by previous surveys.

The results obtained from the present work are compared with previous measurements from the 2dfGRS [35] and 6dFGS [36], WiggleZ [31], VVDS [34], VIPERS [37] and FastSound [38] surveys, as well as the BOSS DR12 completed sample [24].
Not all these measurements perform the anisotropic clustering fit using the AP parameters to extract $f\sigma_8$ and are represented with dashed lines. Table 6.4 summarizes the measurements of $f\sigma_8$ and highlights which one have been marginalized over the AP parameters. Previous works have started to explore the $z > 0.8$ redshift range using a sample of blue and red galaxies for VIPERS at $z = 0.86$ [37] and a sample of ELGs for FastSound at $z = 1.4$ [38] and measured $f\sigma_8$ using the monopole and the quadrupole and they reached a $\sim 25\%$ measurement at a fixed D_A and H. In this paper, we do marginalize over D_A and H and we find a $\sim 18\%$ measurement whereas when setting H and D_A to their fiducial values, we find a $\sim 11\%$ measurement.

It is worth highlighting that if a simultaneous fit of $f\sigma_8$, H and D_A is not performed, the uncertainties in $f\sigma_8$ do not reflect the marginalization and as a consequence the measurement cannot be used to constrain models which explore alternative gravity

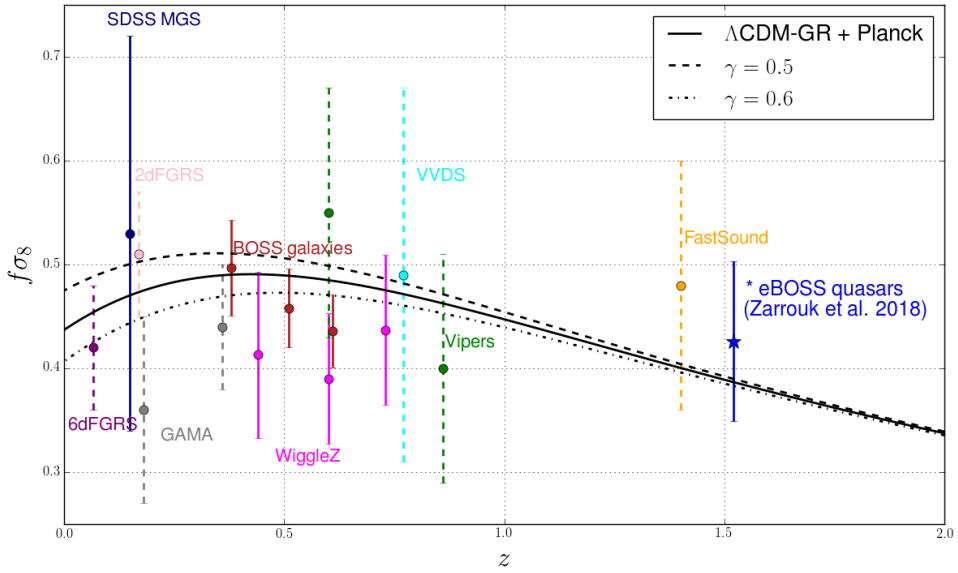


FIGURE 5.13 – Measurements of $f\sigma_8(z)$ with redshift compared to the prediction from the flat Λ-CDM+GR model with *Planck* parameters. The $f\sigma_8(z)$ result presented in this work for the quasar sample is represented by the * marker and is obtained using 3-multipole fit. The error bar represents the total systematic error that includes the statistical precision and the systematic error related to the RSD modeling used in this analysis.

scenario.

5.4.3 Bias measurements

In this kind of clustering analysis, we do not consider the linear bias $b = 1 + F'$ as a cosmological parameter, in particular in this analysis where we showed the two-point correlation function when fitting from $16h^{-1}$Mpc is not sensitive enough to the second bias parameter F'' to constrain it efficiently and varying the prescription on F'' can change the linear bias, although the change represents less than 1σ.
However, several papers have reported a measurement of the quasar linear bias for different samples probing different redshift ranges. It is therefore worth comparing our measurement with theirs. Figure 5.14 shows the measurements of the linear bias as a function of redshift from this analysis for the full redshift range ($z_{\rm eff} = 1.52$), the *lowz* and *highz* bins using eBOSS quasars (in blue), using the first year of eBOSS data [40, 41] (in red and green), using a sample of BOSS quasars [42] (in magenta) and using the 2dF QSO Redshift Survey [43] (in light blue).
The dashed line corresponds to an ad-hoc parametrization

$$b_Q(z) = \left(\frac{1+z}{1+a}\right)^b \tag{5.26}$$

where a and b are obtained by fitting the mentioned measurements which yields $a = 0.46$ and $b = 1.59$.
The dotted line corresponds to the parametrization proposed in [40]

$$b_Q(z) = \alpha[(1+z)^2 - 6.565] + \beta \tag{5.27}$$

with $\alpha = 0.278 \pm 0.018$ and $\beta = 2.393 \pm 0.042$. This parametrization is equivalent to the one proposed in [43]. The bias of quasars is a growing function with redshift, ranging from 1.6 to 3.4 across the eBOSS sample.

We found larger uncertainties in our measurements compared to the one by [40] whereas the size of the sample has more than doubled. In [40], data are fitted from 10 to 85 h^{-1}Mpc using a non linear matter power spectrum P_m obtained from CAMB adding non-linearities and they accounted for linear redshift-space distortions using the Kaiser formula [44] :

$$P_Q(k,\mu) = b_Q^2(1+\beta\mu_k^2)^2 P_m(k) \tag{5.28}$$

where we remind $\beta = f/b_Q$ and they used the GR prediction $f \approx \Omega_m^{0.55}(z)$. The difference could be due to the fact we marginalize over a larger set of parameters, in particular regarding redshift uncertainties where there is a degeneracy between σ_{tot} and F' as shown in figure 5.3. In addition, we already stressed the impact of the second bias parameter prescription on the linear bias.

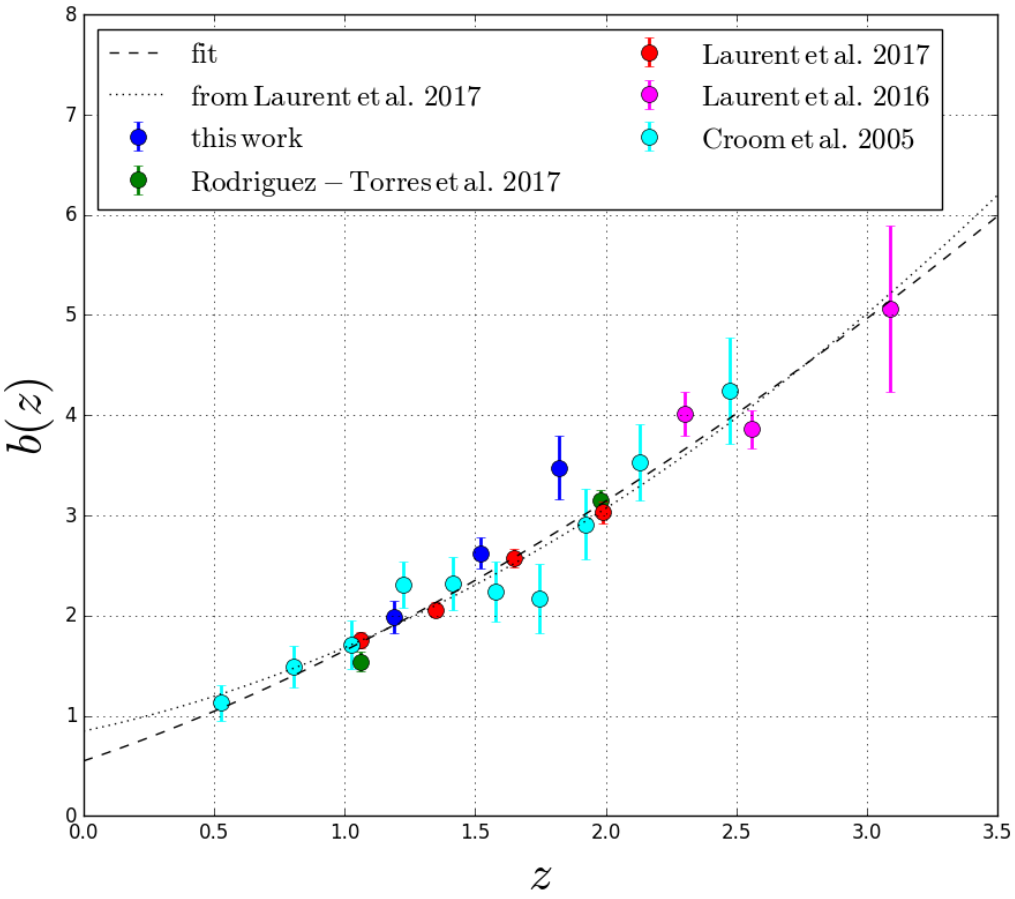


FIGURE 5.14 – Linear bias measurements as a function of redshift for different samples of quasars : the eBOSS DR14 sample (this work, in blue), the first year of eBOSS data taking [40, 41], the BOSS DR12 sample [42] and the 2dFQZ [43]. The dashed curve corresponds to a parametrization from the fit of the mentioned results and the dotted line corresponds to the parametrization in [40].

5.5 Cosmological implications of the eBOSS DR14 quasar measurements

This section aims at discussing the cosmological implications of the eBOSS DR14 quasar sample. As highlighted in the previous sections, our measurements using BAO-only and full-shape analyses are competitive with previous measurements from other surveys but not yet with previous BOSS measurements using galaxies and Ly-α forests. For these reasons, in [2, 1] we did not fully explore the cosmological implications of the eBOSS measurements with the DR14 sample and decided to leave it for the analysis of the final sample whose statistics is expected to double.

Measurements which are independent can be combined by taking the product of their likelihoods. In the case of measurements that are correlated, one needs to take into account the covariance between the measurements as when we combined power spectrum and correlation function results in section 5.3. For each parameter space, we will specify the data sets used to provide constraints.

5.5.1 Constraints on ΛCDM and extensions

In this section, we opt not to include the BAO measurements from other surveys than from BOSS/eBOSS, essentially because the uncertainties are sufficiently large that they would not affect the results but also to avoid volume overlap between different surveys (between WiggleZ and BOSS for instance). BAO measurements can be used to constrain the geometry of the universe through Ω_m and Ω_Λ.
Allowing $\Omega_m + \Omega_\Lambda \neq 1$ can constrain an extension of the ΛCDM model dubbed 'oΛCDM' where the curvature is not fixed. In the ΛCDM model, dark energy is parametrized by a constant Λ whose equation of state $w = p/\rho = -1$. Allowing $w \neq -1$ can constrain an extension of the ΛCDM model dubbed 'wCDM'. Another extension which was mentioned in the introduction consists in allowing the equation of state to vary with time where a standard parametrization is $w(a) = w_0 + w_a(1-a)$ (see section 1.1.5.1 for more details). We do not explore the cosmological implications of the eBOSS DR14 quasar sample in the plane w_0-w_a and leave it for future work with the final eBOSS sample.

5.5.1.1 BAO-only analysis

In [2], we follow the approach adopted in [45] to constrain the geometry of the universe by using the BAO measurements only. To do so, we assume only that the BAO feature has a constant comoving size with no knowledge of the physics that produced this feature. We use an open ΛCDM cosmology which is parametrized using three parameters, Ω_m, Ω_Λ and $H_0 r_d$. We can constrain Ω_m and Ω_Λ, whose sum is related to the cosmic geometry at the present epoch, by performing a MCMC fitting using a modified version of CosmoMC [46] and then by marginalizing over $H_0 r_d$.

Figure 5.15 presents the 68% and 95% confidence level joint constraint in the plane Ω_m-Ω_Λ (left panel) and the one-dimensional probability distribution of Ω_Λ using three data sets which are summarized in table 5.10. Note that the latest cross-correlation results of quasars with Ly-α forests using the DR12 sample [33] was not yet published at the time of writing the eBOSS DR14 quasar BAO paper.

The quasar BAO measurement improves the current constraints on Ω_m and Ω_Λ compared to BOSS DR12 galaxies alone. The existence of dark energy is raised from 2.9σ to 3.4σ confidence-level when the eBOSS quasar BAO is added to BOSS galaxies. This is because it provides a high-redshift constraint, however the tightest constraints mainly arises from

TABLE 5.10 – Data sets used to constrain the Ω_m-Ω_Λ plane using BAO-only measurements in [2].

data sets	description	reference
BOSS galaxies	anisotropic BAO measurement from BOSS DR12 galaxies	[24]
BOSS galaxies+eBOSS	anisotropic BAO measurement from BOSS DR12 galaxies	[24]
	isotropic BAO measurement using eBOSS DR14 quasars	[2]
full BAO	anisotropic BAO measurement from BOSS DR12 galaxies	[24]
	isotropic BAO measurement using eBOSS DR14 quasars	[2]
	anisotropic BAO measurement from the DR12 Ly-α auto-correlation	[25]
	anisotropic BAO measurement from the DR11 Ly-α cross-correlation	[26]
	isotropic BAO measurement from SDSS MGS	[32]
	isotropic 6dFGRS galaxy samples	[36]

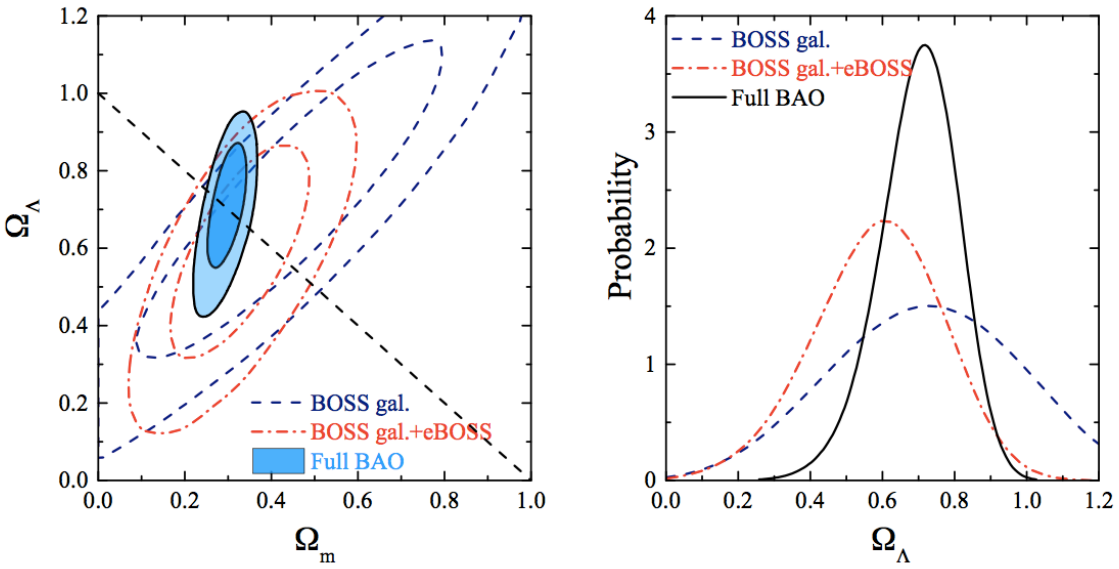


FIGURE 5.15 – Left : The 68% and 95% confidence level contour plots for Ω_m and Ω_Λ using three data sets presented in table 5.10. Here, we assume only that the BAO scale is constant with redshift (see the text for more details). The dashed line illustrates a flat universe with $\Omega_m + \Omega_\Lambda = 1$. Right : The one-dimensional probability distribution of Ω_Λ using the same three data sets. From [2].

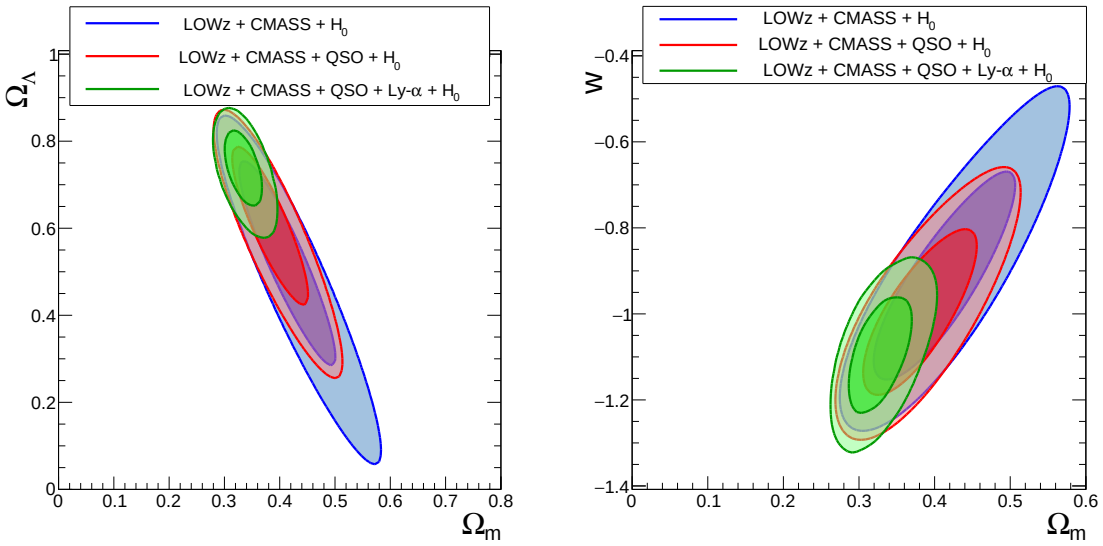


Figure 5.16 – Left : Cosmological constraints in the Ω_Λ vs Ω_m plane. All contours are shown assuming a flat ΛCDM-model. Right : Cosmological constraints in the w vs Ω_m plane. The inner and outer contours show the 68 and 95% confidence-level two-dimensional marginalized constraints. The blue contour represents the cosmological constraints using BOSS DR12 galaxies, the red contour shows the gain when adding the eBOSS quasar sample and the green contour also includes the results from Ly-α measurements although they are not obtained from full-shape analysis.

the Ly-α measurements at redshift 2.4. A 3.4σ detection of dark energy is reached using the BOSS DR12 galaxies and eBOSS DR14 quasars and goes to a 6.6σ detection when considering the 'full BAO' data set, demonstrating the level arm of Ly-α forests studies at high redshift. It also highlights the robustness and constraining power of the BAO-only technique.

5.5.1.2 Full-shape analysis

Similarly to the BAO-only analysis of the eBOSS DR14 quasar sample [2], we evaluate the impact of our distance measurements from the analysis of the full-shape correlation function on extensions of ΛCDM. The left panel (resp. right panel) of figure 5.16 shows the contour in the Ω_Λ vs Ω_m plane (resp. w vs Ω_m) to test predictions of oCDM (resp. wCDM). We see that, when using H_0 from *Planck*, adding the current eBOSS quasar RSD measurement (red contour) to the BOSS DR12 sample [blue contour, 24] substantially improves the constraints on the extensions of ΛCDM. For comparison with figure 5.15, we also add the Ly-α BAO measurements (green contour) even if they do not come from full-shape analysis. As highlighted in the next section, Ly-α measurements provide an additional strong constraint in full agreement with a flat and pure cosmological constant universe.

As explained in section 1.1.2.2, the CMB provides a lot of information about the physics in the early universe. But it is also a powerful probe of the presence of dark energy or a cosmological constant. The late-time modifications of gravity can affect the CMB anisotropies and polarization. We briefly mention some of the most important effects of dark energy or modified gravity models on the CMB anisotropies, further details can be found in [47] :

— change the expansion history and hence the distance to the last scattering surface

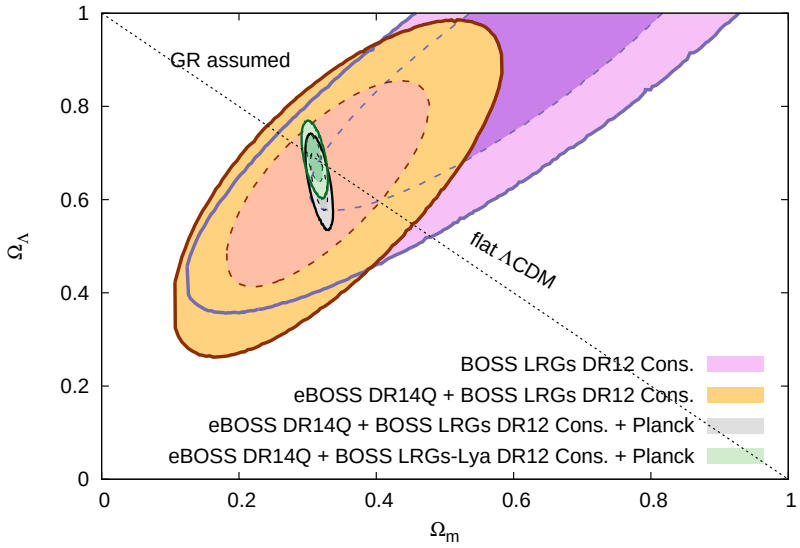


FIGURE 5.17 – The 68% and 95% confidence level contour plots for Ω_m and Ω_Λ for BOSS DR12 galaxies alone (pink), combined with eBOSS DR14 quasar (orange), and when adding *Planck* data (green and grey contours). The combination of CMB and large-scale structure observations is very powerful as each probe is sensitive to cosmology in a different way.

which would lead to a shift in the peaks [48]

— change the growth of structure which would lead to a mismatch between the amplitude of the fluctuations A_s and late-time measurements of σ_8 [49, 50]

— change in the lensing potential due to additional dark energy perturbations or modifications of GR [51, 52]

— affects the ISW effect (low-l region of figure 1.9 by causing the decay of gravitational potential at late times affecting the CMB anisotropies at low l (large angular scales) [53] or by enhancing the cross-correlation between the CMB and LSS [54]

Therefore, by combining CMB constraints from *Planck* [17] with LSS measurements, we can break degeneracies and provide tighter constraints. The grey contour in figure 5.17 shows the constraints on Ω_Λ and Ω_m when adding *Planck* data. It illustrates well the constraining power of combining independent probes.
The combination of BOSS DR12 galaxies, BOSS DR12 Ly-α, eBOSS DR14 quasars and *Planck* yields the following constraints [3] on Ω_m, Ω_Λ and hence the curvature density Ω_k,

$$\Omega_m = 0.3094^{+0.0076}_{-0.0080} \tag{5.29}$$

$$\Omega_\Lambda = 0.697^{+0.035}_{-0.032} \tag{5.30}$$

$$\Omega_k = -0.007 \pm 0.030 \tag{5.31}$$

The results are in agreement with an universe dominated by dark energy with no curvature.

5.5.2 Constraining alternative gravity models

In section 1.1.5, we reviewed some examples of alternative gravity models. In general, the modifications of GR make an order-unity change in the dynamics at cosmological scales. At the solar-system scales, these modifications need to have a very small or negligible effect and it is usually the case by invoking non-linear screening mechanisms, such as the Chameleon or the Vainshtein effects, to restore GR in high density regions and agree with the local tests of GR.

First detection of a gravitational wave associated with a gamma-ray burst

In September 2015, the first detection of gravitational waves from a black hole merger was confirmed using the Laser Interferometer Gravitational-Wave Observatory [LIGO, 55]. In August 2017 it was followed by the detection of gravitational waves from the merger of two neutron stars using the two interferometers of LIGO. The recent addition of the interferometer Virgo allowed to localize the source by the absence of signal from Ligo which suggested that the source was located in the parallel direction of the mirrors [56] (we refer to this event as GW170817). Thanks to the localization of the source, it was possible to relate the emission of gravitational waves with the emission of a short burst of gamma-rays from a similar location in the sky few seconds after using the Fermi satellite [57]. Follow-up observations by other telescopes confirmed that the gravitational wave and gamma rays came from the same source, a binary neutron star merger in the NGC 4993 galaxy at $z = 0.009787$, which is approximately 130 million light years away from Earth.
The association of both events allowed to make a very precise measurement of the speed of gravitational waves. It is compatible with the speed of light with a time delay [58],

$$\Delta t = \int_{a_e}^{1} \frac{da}{aH}\left(1 - \frac{c}{c_g(a)}\right) + \delta t = 1.74 \pm 0.05 s \tag{5.32}$$

where a_e is the scale factor associated with the redshift of the host galaxy, $z_e = 0.009787$ and δt is the time delay between the gravitational wave and light emission where [58] provide a conservative constraint $\delta t \in [-1000s, 100s]$. The fact that the two signals traveled from this distance to Earth with just few seconds delay implies that gravitational waves travel at the same speed as light to within $c_g/c - 1 < 10^{-15}$. Previous constraints on the relative speeds had only been at the level of 10^{-5}, so this single observation improved our knowledge of a fundamental property of gravity by 10 orders of magnitude. Using this recent detection, four research groups have now placed some of the tightest constraints to date on modified gravity scenarios which predict different speeds for gravitational waves compared to light [59, 60, 61, 62]. Indeed, additional fields coupled to gravity can affect the propagation speed of graviatational waves, this is the case for the full Galileon and for most of the scalar-tensor theories whose equations of motion are at second order in time derivatives at most. Among the viable theories after GW170817, quintessence and $f(R)$ theory survive for instance.

A general way to parametrize modified gravity theories is to specify the relation between the two gravitational potentials Ψ and Φ which govern the motion of matter and of light respectively. Indeed, modified gravity models affect the clustering of galaxies and changes how mass affects the propagation of light. As a consequence the Poisson equations are modified accordingly and a way of parametrizing these modifications is to introduce two dimensionless parameters G_M and G_L [e.g. 63] :

$$\nabla^2\Psi = 4\pi \mathrm{G}_N a^2 \rho \Delta \mathrm{G}_\mathrm{M} \tag{5.33}$$

$$\nabla^2(\Psi + \Phi) = 8\pi \mathrm{G}_N a^2 \rho \Delta \mathrm{G}_\mathrm{L} \tag{5.34}$$

where G_N is the Newton constant.
Alternatively one could also use the ratio of the two potentials, referred to as the gravitational slip γ_{slip} instead of $\mathrm{G_L}$ to parametrize the modified Poisson equations. In the general case, G_M and G_L are functions of time and scale. Therefore, special care is required when comparing theoretical predictions with measurements which are obtained assuming a scale-independent growth rate over the fitting range.

5.5.2.1 Constraints on the growth index

Allowing the growth index γ to vary is one of the minimal extension to GR that has been widely used to search for any departure from its GR+ΛCDM prediction. We remind the prediction for the growth rate given by [64, 65] :

$$f(z) = [\Omega_m(z)]^\gamma \text{ where } \gamma = \frac{3(1 - w_{\mathrm{DE}})}{5 - 6w_{\mathrm{DE}}} \tag{5.35}$$

where $w_{\mathrm{DE}} = -1$ for a cosmological constant which gives $\gamma = 0.55$ in the ΛCDM model. The ΛCDM+GR prediction cannot be accurately tested given the statistical precision of the eBOSS quasar sample only. Combining our data to the measurement of Ω_m from *Planck* produces $\gamma = -0.2 \pm 1.2$. The lack of precision also arises because in the eBOSS quasar redshift range, Ω_m is close to 1 and the sensitivity to γ is therefore reduced as can be seen from the black curves in figure 6.9, which shows theoretical predictions of $f\sigma_8$ for different values of γ. The lack of sensitivity is also shown by the pink contour in the top panel of figure 5.18 which represents the eBOSS DR14 quasars + *Planck* constraints on the Ω_m-γ plane. The combination of *Planck* with BOSS DR12 galaxies and eBOSS DR14 quasars is shown by the grey contour. Table 5.11 summarizes the most recent measurements of γ using the BOSS DR12 galaxies alone with *Planck* and the BOSS DR12 galaxies with the eBOSS DR14 quasar sample which can be compared directly. A recent paper [66] explored the cosmological implications of the latest observations in the context of modified gravity models and we reported their constraints on γ for two data sets :

— 'ALL17' which includes LSS measurements (BOSS DR12 galaxies BAO and RSD and other BAO measurements), CMB measurements using *Planck* data, SN JLA sample, H_0 measurments using local probes and the ages of passive galaxies and weak lensing measurement using the CFHTLens sample

— 'ALL18' which includes 'ALL17' with tomographic BAO and RSD measurements using the eBOSS DR14 quasar sample in [12]. These measurements account for the redshift evolution of the cosmological parameters better compared to measurements at effective redshifts. However, at the level of the current precision of the DR14 quasar sample, no significant improvement has been found and all the full-shape measurements are comparable (more details can be found in section 5.2.)

Table 5.11 – Constraints on the growth index from various data sets.

data sets	reference	γ
Planck+BOSS DR12 galaxies FS	[67]	0.52 ± 0.10
Planck+BOSS DR12 galaxies CF	[68]	0.609 ± 0.079
Planck+BOSS DR12 galaxies+eBOSS DR14 quasars	[3]	0.55 ± 0.19
ALL17	[66]	0.506 ± 0.031
ALL18	[66]	0.485 ± 0.031

The bottom panel of figure 5.18 shows the 68% and 95% confidence level contour for γ and σ_8 derived in [66]. We focus the discussion on the comparison between the red contour ('ALL17') and the orange contour ('ALL18'). The contour is slightly shifted but the surface is unchanged, meaning that for the moment the constraining power of the eBOSS quasar sample is limited in the plane γ - σ_8.

5.5.2.2 Galileon theories

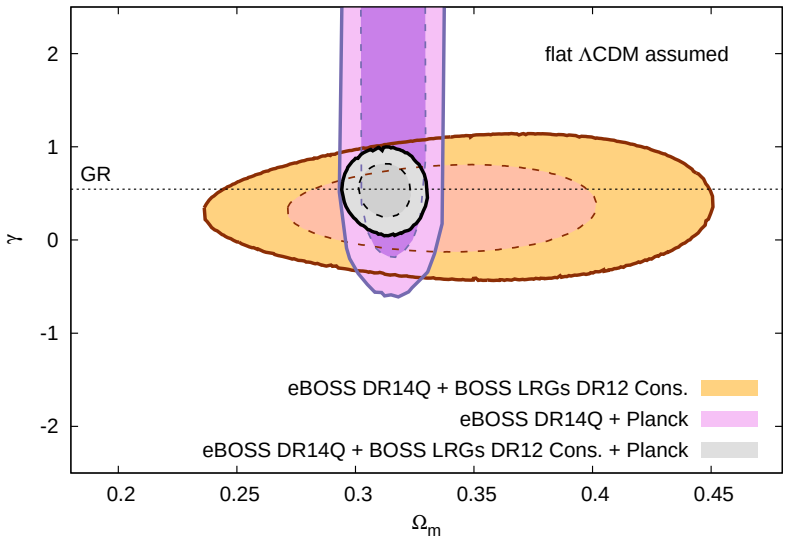
We remind the action of the disformally coupled Galileon in the Einstein frame given by equation 1.67 in section 1.1.5.2 :

$$S_{\text{Galileon}} = \int d^4x\sqrt{-\tilde{g}}\left[\frac{M_P^2}{2}R - \frac{1}{2}\sum_{i=1}^{5}\frac{c_i}{M^{3(i-2)}}L_i - \frac{M_P}{M^3}c_G\partial_\mu\partial_\nu T^{\mu\nu}\right] \tag{5.36}$$

where $L_1 = \pi$ behaves like a cosmological constant, so $c_1 = 0$ is imposed as we look for alternatives to the cosmological constant and c_G is the disformal coupling to matter as the conformal coupling is disfavored by observations [69]. The presence of couplings with matter can be interpreted again as a fifth force while the presence of non-linear Lagrangians ($L_{3,4,5}$) are necessary to screen this fifth force at small scales, in the vicinity of massive objects using the Vainshtein effect [70]. We call the 'cubic Galileon' the Galileon model up to L_3 and 'full Galileon' the Galileon model with the five Lagrangians.

Leloup et al. (in prep) derive the Galileon predictions using their own modified version of CAMB and performed an MCMC exploration of the parameter space using their own modified version of CosmoMC using different data sets : CMB temperature, polarization and lensing maps using *Planck* data, BAO constraints using BOSS DR12, WiggleZ, SDSS MGS and 6dFGS and JLA SN sample. Then, they perform a posteriori comparison to the speed of the gravitational wave from GW170817 event for the full Galileon where the speed of gravitational wave depends on $c_{4,5,G}$ (while for the cubic Galileon, the speed of gravitational wave is equal to the speed of light). They show that even without considering the GW constraint, the disformally coupled Galileon model is very disfavored by current data. In addition, the time delay between gravitational wave and light they infer for the full Galileon model from their best-fit values is between 10 and 100 years which is much higher than the constraint put by GW170817 (see the box above). Therefore, the full Galileon can be considered as ruled out.

However, since Leloup et al. (in prep) did not consider the growth rate measurements as an additional probe, we just show a posteriori comparison between Galileon predictions for $f\sigma_8$ and the latest observations in figure 5.19. The predictions for $f\sigma_8$ using the best-fitting values for the cubic and full Galileon models of Leloup et al. (in prep) are shown in dashed and solid red lines respectively, while the prediction from the uncoupled Galileon model of [69] is shown in dotted red line.



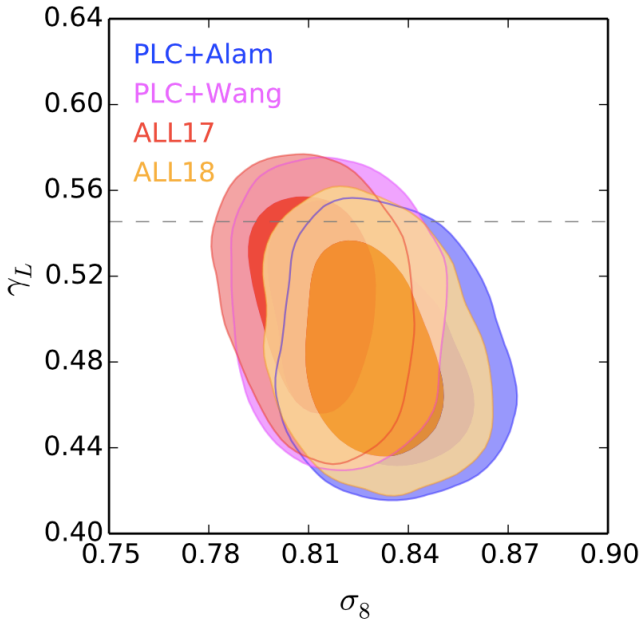


FIGURE 5.18 – Top panel : The 68% and 95% confidence level contour plots for Ω_m and γ using DR14 eBOSS quasars with *Planck* (pink), DR14 eBOSS quasars with BOSS DR12 galaxies (orange), DR14 eBOSS quasars + BOSS DR12 galaxies + *Planck* (grey).The dotted line shows the ΛCDM+GR prediction : $\gamma = 0.55$. From [3]. Right panel : The 68% and 95% confidence level contour plots for γ and σ_8 using different data sets presented in table 5.11. From [66].

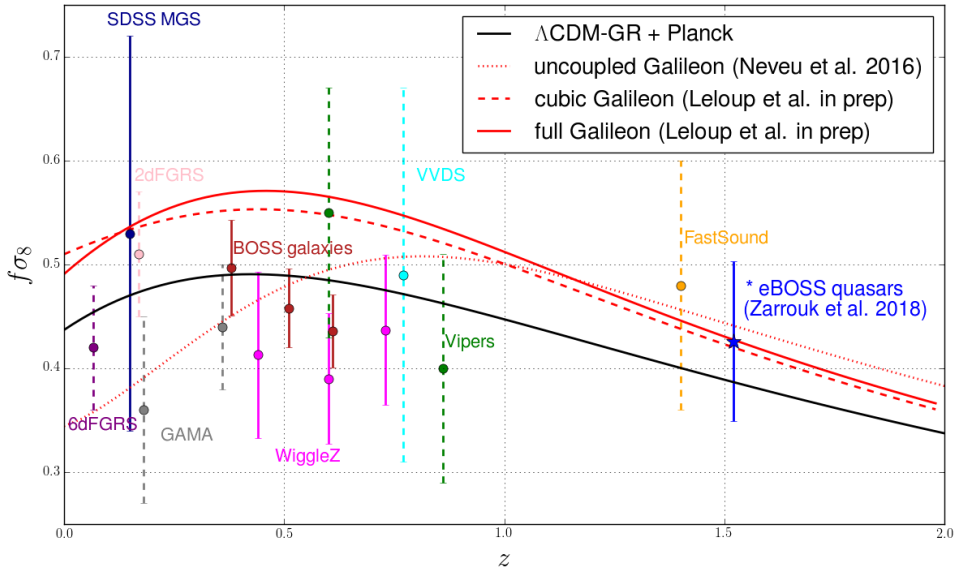


FIGURE 5.19 – Growth rate measurements compared to Galileon predictions for the full disformally coupled case (solid red), the cubic disformally coupled case (dashed red) and the full uncoupled case (dotted red).

We could be tempted to interpret the figure saying that measurements of the growth rate at $z \sim 0.5$ are in $\sim 3\sigma$ tension with the predictions of the full and cubic Galileon theories. However, the comparison between predictions of $f\sigma_8$ from modified gravity models and measurements is not trivial because of two effects at least :

— **Scale-dependent growth rate** : The prediction for the growth rate can depend on the scale which is considered as we mentioned previously. Fortunately, this is not the case for Galileon theories which introduce a scale-independent effective gravitational coupling G_{eff} in the modified Poisson equations as follows :

$$\nabla^2\Psi = 4\pi G_{\text{eff}} a^2 \rho \Delta \tag{5.37}$$

— **Non-linearities** : The Galileon predictions for the growth rate are obtained solving the equation that governs the growth function in the linear regime. Linear perturbation theory in the ΛCDM+GR framework breaks down on scales $k > 0.05\,h\text{Mpc}^{-1}$, taking into account theoretical uncertainties could go up to $k_{max} = 0.10 - 0.15\,h\text{Mpc}^{-1}$ depending on redshift. However, current analyses with modeling of the non-linear power spectrum can go up to $k = 0.3\,h\text{Mpc}^{-1}$. Some studies used N-body simulations in a Galileon theory to investigate the effect of non-linearities in the matter power spectrum prediction [71, 72]. In particular, [71] found that for k-scales between 0.1 and $0.4\,h\text{Mpc}^{-1}$, non-linearities affect the matter power spectrum by 5% only at redshifts $z < 0.2$. More important deviations and unphysical solutions in dense regions have been reported in [72] (resp.[73]) when they studied the quartic Galileon (resp. full Galileon) models, making the comparison with observations more difficult.

5.5.2.3 Scalar-tensor theories with chameleon effect

In this section, we comment the results obtained by [66] who used another widely used parametrization of the potentials based on $\mu(a,k)$ and $\eta(a,k)$ which are functions of the scale factor (hence redshift) and the scale (here in k-space) that modify the Poisson equations as follows :

$$k^2\Psi = -4\pi G_N a^2 \mu(a,k)\rho\Delta \tag{5.38}$$

$$\frac{\Phi}{\Psi} = \eta(a,k) \tag{5.39}$$

where in GR we have $\Psi = \Phi$ and $\mu(a,k) = \eta(a,k) = 1$. Here, the growth of structure depends on the scale, and we follow the approach in [66] that treat the scale-independent and scale-dependent cases separately.

Scale-independent case A general parametrization for $\mu(a)$ and $\eta(a)$ is to use power-law functions [74] :

$$\mu(a) = 1 + \mu_s a^s \tag{5.40}$$

$$\eta(a) = 1 + \eta_s a^s \tag{5.41}$$

$$\tag{5.42}$$

where we report the results for $s = 1$ and $s = 3$ found by [66] in figure 5.20. The comparison between 'ALL17' (red) and 'ALL18' (orange) whose only difference is the inclusion of the RSD measurements using the eBOSS DR14 quasar sample shows that the additional gain provided by the DR14 sample is marginal. The associated mean and 68% confidence-level uncertainties on μ and η are :

data set	s	$\mu - 1$	$\eta - 1$
ALL17	1	-0.131 ± 0.075	0.863 ± 0.286
	3	-0.405 ± 0.184	2.555 ± 0.835
ALL18	1	-0.132 ± 0.075	0.873 ± 0.289
	3	-0.398 ± 0.184	2.516 ± 0.832

The combination of all these data sets excludes the case ($\mu_s = \eta_s = 0$) at 2.2σ and 3.1σ levels for $s = 1$ and $s = 3$ respectively. It can be compared to the GR prediction where $\mu = \eta = 1$ (i.e. $\mu_s = \eta_s = 0$) which therefore seems to be strongly disfavored by data. However, the parametrization given by equations 5.42 adds two parameters compared to GR, so one should compute the Bayesian evidence to compare models with different parameter spaces. Moreover, additionnal data is definitely necessary before setting any conclusion.

Scale-dependent case For general scalar-tensor theories, $\mu(a,k)$ and $\eta(a,k)$ can be parametrized and is referred to as the BZ parametrization [75] :

$$\mu(a,k) = \frac{1 + \beta_1\lambda_1^2 k^2 a^s}{1 + \lambda_1^2 k^2 a^s} \tag{5.43}$$

$$\eta(a,k) = \frac{1 + \beta_2\lambda_2^2 k^2 a^s}{1 + \lambda_2^2 k^2 a^s} \tag{5.44}$$

where $\beta_{1,2}$ are dimensionless couplings, s the power index and $\lambda_{1,2}$ are length scales.

$f(R)$ **theories** are a special case of scalar theories whose action is given by :

$$S_{\mathrm{f(R)}} = \int d^4x\sqrt{-g}\frac{-M_P^2}{2}f(R) \tag{5.45}$$

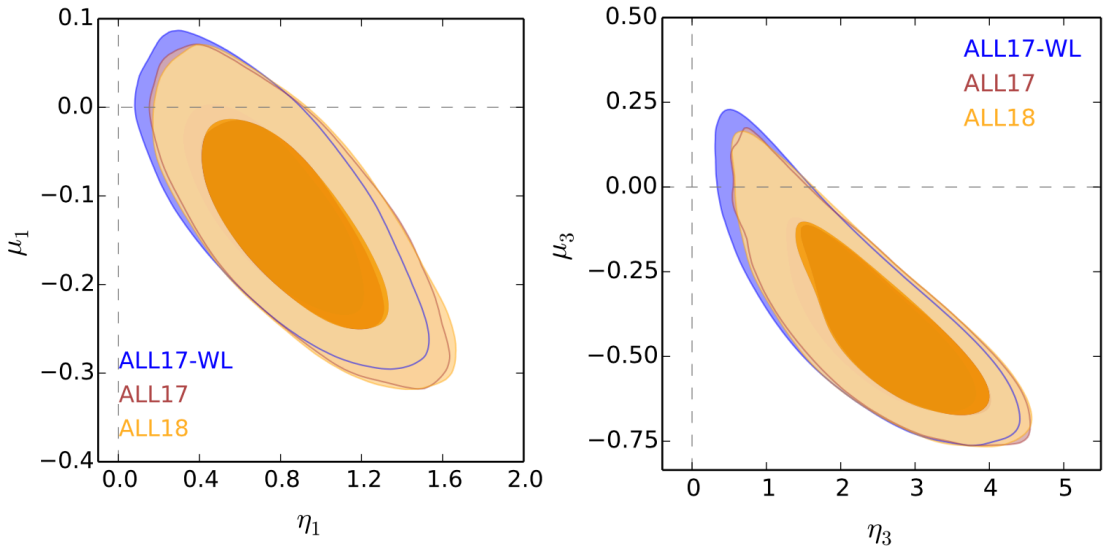


FIGURE 5.20 – The 68% and 95% contour plots for μ_s and η_s where the left panel is for $s = 1$ and the right panel for $s = 3$. We focus on the difference between ALL17 (red) and ALL18 (orange) when adding the RSD measurement using the eBOSS DR14 quasar sample.

We usually work with a simple version of $f(R)$ in which the background evolution (homogeneous and isotropic metric) is set to match that of ΛCDM. In this case, the additional degree of freedom is encoded in only one parameter by fixing $\beta_1 = 4/3$, $\beta_2 = 1/2$ and $\lambda_2^2/\lambda_1^2 = 4/3$. In [66] they fix $s = 4$ following [76] to closely reproduce the ΛCDM model for the background evolution while scalar perturbations of the metric lead to the modified Poisson equations defined by equation 5.39. Therefore, there is only one free parameter, λ_1 but usually constraints are put on $\log B_0$ where $B_0 = 2H_0^2\lambda_1^2/c^2$ is dimensionless. The GR limit corresponds to $B_0 = 0$. The left panel of figure 5.21 shows the 68% and 95% confidence-level contour plots for Ω_m and $\log B_0$. The blue contour uses BOSS DR12 BAO and RSD results at three effective redshifts while the pink contour uses the tomographic BAO and RSD results on the same data sample (i.e. BOSS DR12 galaxies) performed by [77] using a similar technique to the hybrid redshift-weighting technique in [12] that we presented in section 5.2.1.2. We can see the improvement between the two contours when taking into account the redshift evolution of the cosmological parameters. The improvement is not that significant when adding the tomographic RSD measurements of [12]. The final upper limit derived on $\log B_0$ using ALL18 is $\log B_0 < -4.93$ at 95% confidence-level, which represents an improvement compared to the latest constraint $\log B_0 < -4.54$, set by [18] using several data sets among which the consensus BAO and RSD measurements of [24].

For **general scalar-tensor theories**, the following consistency relations hold : $\beta_1 = \lambda_1^2/\lambda_2^2$ and $\beta_2 = 2/\beta_1 - 1$. These relations could reduce the parameter space but [66] did not apply them as a constraint and rather perform a direct comparison a posteriori with their best-fitting constraint. The right panel of figure 5.21 shows the 68% and 95% confidence level contour plots for β_1 and β_2 derived from the results using ALL18 for two priors on s : $s \in [0, 10]$ (grey contour) and $s \in [1, 4]$ (orange contour). It yields the following constraints :

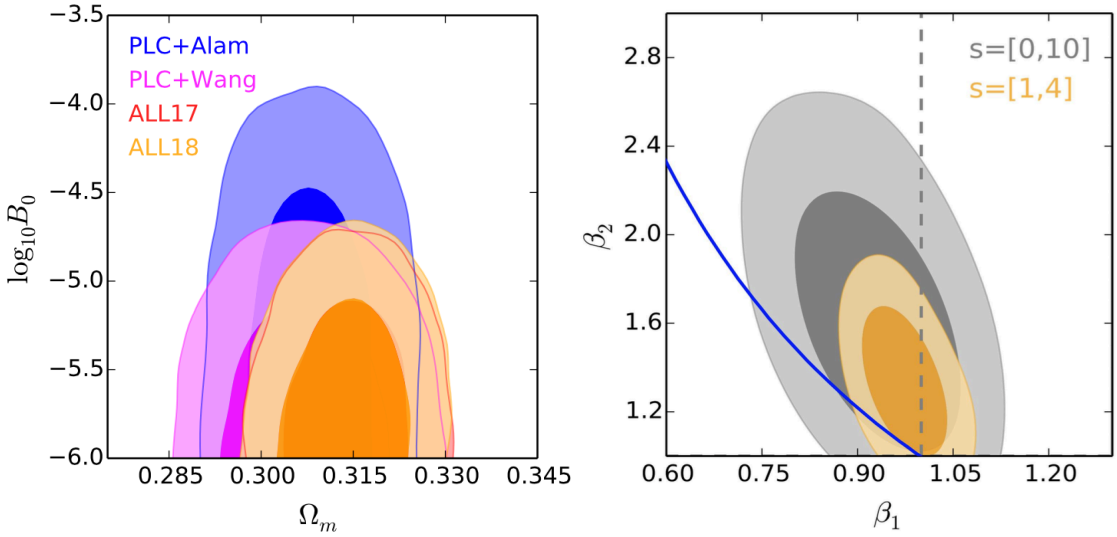


FIGURE 5.21 – The 68% and 95% contour plots for $f(R)$ theory in the plane Ω_m - $\log B_0$ for different data sets (left panel) and for more general scalar-theories in the plane β_1 - β_2 for 'ALL18' using two priors on s.

parameter	$s \in [1,4]$	$s \in [0,10]$
β_1	0.974 ± 0.033	0.928 ± 0.061
β_2	1.349 ± 0.165	1.647 ± 0.296

According to figure 5.21, the special $f(R)$ which mimics the ΛCDM background evolution with $\beta_1 = 4/3$, $\beta_2 = 1/2$ and $s = 4$ is strongly disfavored by data compared to general scalar-tensor models following the BZ parametrization. Thanks to the upcoming data, we expect to provide tighter constraints, and even to rule out some $f(R)$ models more strongly.

Growth of structure Figure 5.22 shows the predictions for the $f\sigma_8$ for the $f(R)$ model favored by 'ALL18' along with the predictions for the general case with $\beta_{1,2}$ favored by 'ALL18' and the scale-independent case with power laws $s = 1, 3$. The lever arm to discriminate between these models seems more interesting at low redshifts ($z < 0.5$). The Bright Galaxy Sample (BGS) at median redshift 0.2 that DESI will provide would be of particular interest to constrain the scalar-tensor theories following the BZ form (equation 5.44).

However, the comparison between the predictions of these models and measurements are again not trivial, for the same reasons as for the Galileon : scale-dependent growth rate and non-linear effects. $f(R)$ models have been one of the most studied modified gravity theories and although the results are not necessarily valid for other scalar theories, it is interesting to have in mind the order of magnitude of both effects :

— **Scale-dependence** : Compared to the Galileon, these scalar-theories predict a scale-dependent growth rate. Figure 5.23 shows the f(R) prediction for the growth rate f as a function of k-scale for the f(R) favored by 'ALL18' with $\log B_0 = -4.392$ for three redshifts $z = 0.5, 1.0, 1.5$ (red,green,blue). The scale dependence is small with a ∼2% effect up to $k = 0.2h\text{Mpc}^{-1}$ at $z = 0.5$ where the dependence is more important. It seems to be negligible for this f(R) theory given the present precision of the data, but it is not necessarily the same for all the scalar-tensor theories.

— **Non-linearities** : Using N-body simulations in the f(R) framework, [78] looked at

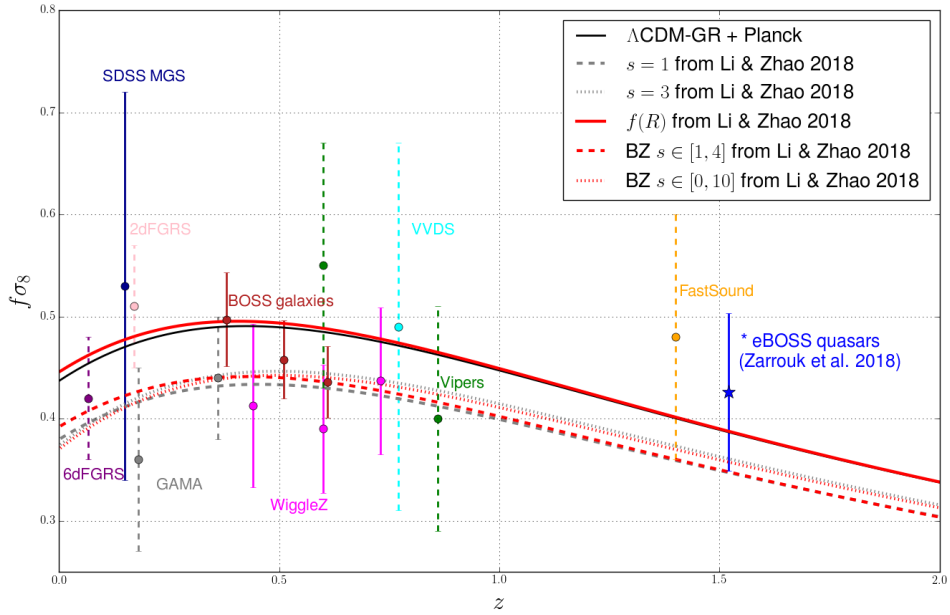


FIGURE 5.22 – Predictions of $f\sigma_8$ as a function of redshift from $f(R)$ theory with $\log B_0 = -4.392$, general scalar-tensor theories in the BZ form with the two priors on the power index s and the two scale-independent predictions for power laws with $s = 1, 3$. All predictions correspond to models favored by 'ALL18'. The prediction from ΛCDM+GR is also shown in black, along with the measurements of $f\sigma_8$.

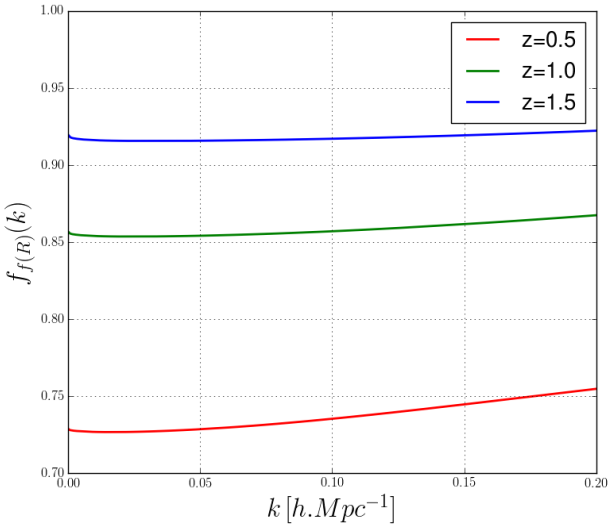


FIGURE 5.23 – The growth rate prediction from $f(R)$ theory with $\log B_0 = -4.392$ favored by 'ALL18' [66] as a function of k-scale for three redshifts : $z = 0.5, 1, 1.5$.

the difference in matter power spectrum between f(R) models with predictions from linear perturbation theory and when including non-linear effects. At $z = 0$ where non-linearities are the most important and for $\log B_0 < -5$, they found a ∼6% difference for $k < 0.3h\text{Mpc}^{-1}$.

Bibliographie

[1] P. Zarrouk, E. Burtin, H. Gil-Marín, A. J. Ross, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2*, **477**, 1639 (2018), doi:10.1093/mnras/sty506.

[2] M. Ata, F. Baumgarten, J. Bautista, F. Beutler, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : First measurement of Baryon Acoustic Oscillations between redshift 0.8 and 2.2*, ArXiv e-prints (2017), arXiv:1705.06373.

[3] H. Gil-Marín, J. Guy, P. Zarrouk, E. Burtin, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : structure growth rate measurement from the anisotropic quasar power spectrum in the redshift range 0.8 z 2.2*, **477**, 1604 (2018), doi:10.1093/mnras/sty453, arXiv:1801.02689.

[4] F. Zhu, N. Padmanabhan, et M. White, *Optimal redshift weighting for baryon acoustic oscillations*, **451**, 236 (2015), doi:10.1093/mnras/stv964, arXiv:1411.1424.

[5] R. Ruggeri, W. J. Percival, H. Gil-Marín, F. Zhu, *et al.*, *Optimal redshift weighting for redshift-space distortions*, **464**, 2698 (2017), doi:10.1093/mnras/stw2422, arXiv:1602.05195.

[6] F. Zhu, N. Padmanabhan, M. White, A. J. Ross, *et al.*, *Redshift weights for baryon acoustic oscillations : application to mock galaxy catalogues*, **461**, 2867 (2016), doi:10.1093/mnras/stw1515, arXiv:1604.01050.

[7] R. Ruggeri, W. J. Percival, E.-M. Mueller, H. Gil-Marin, *et al.*, *The extended Baryon Oscillation Spectroscopic Survey (eBOSS) : testing a new approach to measure the evolution of the structure growth*, ArXiv e-prints (2017), arXiv:1712.03997.

[8] J. Hou, A. G. Sánchez, R. Scoccimarro, S. Salazar-Albornoz, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : anisotropic clustering analysis in configuration-space*, ArXiv e-prints (2018), arXiv:1801.02656.

[9] R. Ruggeri, W. J. Percival, H. Gil Marin, F. Beutler, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : measuring the evolution of the growth rate using redshift space distortions between redshift 0.8 and 2.2*, ArXiv e-prints (2018), arXiv:1801.02891.

[10] G.-B. Zhao, Y. Wang, S. Saito, H. Gil-Marín, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : a tomographic measurement of cosmic structure growth and expansion rate based on optimal redshift weights*, ArXiv e-prints (2018), arXiv:1801.03043.

[11] F. Zhu, N. Padmanabhan, A. J. Ross, M. White, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : Measuring the anisotropic Baryon Acoustic Oscillations with redshift weights*, ArXiv e-prints (2018), arXiv:1801.03038.

[12] D. Wang, G.-B. Zhao, Y. Wang, W. J. Percival, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : anisotropic Baryon Acoustic Oscillations measurements in Fourier-space with optimal redshift weights*, **477**, 1528 (2018), doi:10.1093/mnras/sty654, arXiv:1801.03077.

[13] H. Gil-Marín, C. Wagner, L. Verde, C. Porciani, *et al.*, *Perturbation theory approach for the power spectrum : from dark matter in real space to massive haloes in redshift space*, **11**, 029 (2012), doi:10.1088/1475-7516/2012/11/029, arXiv:1209.3771 [astro-ph.CO].

[14] R. Scoccimarro, *Redshift-space distortions, pairwise velocities, and nonlinearities*, **70**, 083007 (2004), doi:10.1103/PhysRevD.70.083007, astro-ph/0407214.

[15] A. Taruya, T. Nishimichi, et S. Saito, *Baryon acoustic oscillations in 2D : Modeling redshift-space power spectrum from perturbation theory*, **82**, 063522 (2010), doi:10.1103/PhysRevD.82.063522, arXiv:1006.0699.

[16] K. C. Chan et R. Scoccimarro, *Halo sampling, local bias, and loop corrections*, **86**, 103519 (2012), doi:10.1103/PhysRevD.86.103519, arXiv:1204.5770 [astro-ph.CO].

[17] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, *et al.*, *Planck 2015 results. XIII. Cosmological parameters*, **594**, A13 (2016), doi:10.1051/0004-6361/201525830, arXiv:1502.01589.

[18] E.-M. Mueller, W. J. Percival, et R. Ruggeri, *Optimising primordial non-Gaussianity measurements from galaxy surveys*, ArXiv e-prints (2017), arXiv:1702.05088.

[19] M. Tegmark, A. N. Taylor, et A. F. Heavens, *Karhunen-Loève Eigenvalue Problems in Cosmology : How Should We Tackle Large Data Sets ?*, **480**, 22 (1997), doi:10.1086/303939, astro-ph/9603021.

[20] L. Anderson, É. Aubourg, S. Bailey, F. Beutler, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples*, **441**, 24 (2014), doi:10.1093/mnras/stu523, arXiv:1312.4877.

[21] L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample*, **427**, 3435 (2012), doi:10.1111/j.1365-2966.2012.22066.x, arXiv:1203.6594.

[22] D. J. Eisenstein et W. Hu, *Baryonic Features in the Matter Transfer Function*, **496**, 605 (1998), doi:10.1086/305424.

[23] X. Xu, N. Padmanabhan, D. J. Eisenstein, K. T. Mehta, *et al.*, *A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - II. Fitting techniques*, **427**, 2146 (2012), doi:10.1111/j.1365-2966.2012.21573.x, arXiv:1202.0091.

[24] S. Alam, M. Ata, S. Bailey, F. Beutler, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : cosmological analysis of the DR12 galaxy sample*, **470**, 2617 (2017), doi:10.1093/mnras/stx721, arXiv:1607.03155.

[25] J. E. Bautista, N. G. Busca, J. Guy, J. Rich, *et al.*, *Measurement of baryon acoustic oscillation correlations at z = 2.3 with SDSS DR12 Lyα-Forests*, **603**, A12 (2017), doi:10.1051/0004-6361/201730533, arXiv:1702.00176.

[26] A. Font-Ribera, D. Kirkby, N. Busca, J. Miralda-Escudé, *et al.*, *Quasar-Lyman α forest cross-correlation from BOSS DR11 : Baryon Acoustic Oscillations*, **5**, 027 (2014), doi:10.1088/1475-7516/2014/05/027, arXiv:1311.1767.

[27] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, *et al.*, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies*, **465**, 1757 (2017), doi:10.1093/mnras/stw2679, arXiv:1606.00439.

[28] J. E. Bautista, M. Vargas-Magaña, K. S. Dawson, W. J. Percival, *et al.*, *The SDSS-IV extended Baryon Oscillation Spectroscopic Survey : Baryon Acoustic Oscillations at redshift of 0.72 with the DR14 Luminous Red Galaxy Sample*, ArXiv e-prints (2017), arXiv:1712.08064.

[29] F. Beutler, C. Blake, M. Colless, D. H. Jones, *et al.*, *The 6dF Galaxy Survey : baryon acoustic oscillations and the local Hubble constant*, **416**, 3017 (2011), doi:10.1111/j.1365-2966.2011.19250.x, arXiv:1106.3366.

[30] E. A. Kazin, J. Koda, C. Blake, N. Padmanabhan, *et al.*, *The WiggleZ Dark Energy Survey : improved distance measurements to z = 1 with reconstruction of the baryonic acoustic feature*, **441**, 3524 (2014), doi:10.1093/mnras/stu778, arXiv:1401.0358.

[31] C. Blake, S. Brough, M. Colless, C. Contreras, *et al.*, *The WiggleZ Dark Energy Survey : joint measurements of the expansion and growth history at z 1*, **425**, 405 (2012), doi:10.1111/j.1365-2966.2012.21473.x, arXiv:1204.3674.

[32] A. J. Ross, W. J. Percival, et M. Manera, *The information content of anisotropic Baryon Acoustic Oscillation scale measurements*, **451**, 1331 (2015), doi:10.1093/mnras/stv966, arXiv:1501.05571.

[33] H. du Mas des Bourboux, J.-M. Le Goff, M. Blomqvist, N. G. Busca, *et al.*, *Baryon acoustic oscillations from the complete SDSS-III Lyα-quasar cross-correlation function at z* = 2.4, ArXiv e-prints (2017), arXiv:1708.02225.

[34] L. Guzzo, M. Pierleoni, B. Meneux, E. Branchini, *et al.*, *A test of the nature of cosmic acceleration using galaxy redshift distortions*, **451**, 541 (2008), doi:10.1038/nature06555, arXiv:0802.1944.

[35] W. J. Percival, D. Burkey, A. Heavens, A. Taylor, *et al.*, *The 2dF Galaxy Redshift Survey : spherical harmonics analysis of fluctuations in the final catalogue*, **353**, 1201 (2004), doi:10.1111/j.1365-2966.2004.08146.x.

[36] F. Beutler, C. Blake, M. Colless, D. H. Jones, *et al.*, *The 6dF Galaxy Survey : z≈ 0 measurements of the growth rate and σ₈*, **423**, 3430 (2012), doi:10.1111/j.1365-2966.2012.21136.x, arXiv:1204.4725.

[37] A. Pezzotta, S. de la Torre, J. Bel, B. R. Granett, *et al.*, *The VIMOS Public Extragalactic Redshift Survey (VIPERS). The growth of structure at 0.5 < z < 1.2 from redshift-space distortions in the clustering of the PDR-2 final sample*, **604**, A33 (2017), doi:10.1051/0004-6361/201630295.

[38] T. Okumura, C. Hikage, T. Totani, M. Tonegawa, *et al.*, *The Subaru FMOS galaxy redshift survey (FastSound). IV. New constraint on gravity theory from redshift space distortions at z*
1.4, **68**, 38 (2016), doi:10.1093/pasj/psw029, arXiv:1511.08083.

[39] C. Howlett, A. J. Ross, L. Samushia, W. J. Percival, *et al.*, *The clustering of the SDSS main galaxy sample - II. Mock galaxy catalogues and a measurement of the growth of structure from redshift space distortions at z = 0.15*, **449**, 848 (2015), doi:10.1093/mnras/stu2693, arXiv:1409.3238.

[40] P. Laurent, S. Eftekharzadeh, J.-M. Le Goff, A. Myers, *et al.*, *Clustering of quasars in SDSS-IV eBOSS : study of potential systematics and bias determination*, **7**, 017 (2017), doi:10.1088/1475-7516/2017/07/017, arXiv:1705.04718.

[41] S. A. Rodríguez-Torres, J. Comparat, F. Prada, G. Yepes, *et al.*, *Clustering of quasars in the first year of the SDSS-IV eBOSS survey : interpretation and halo occupation distribution*, **468**, 728 (2017), doi:10.1093/mnras/stx454, arXiv:1612.06918.

[42] P. Laurent, J.-M. Le Goff, E. Burtin, J.-C. Hamilton, *et al.*, *A 14 h^{-3} Gpc^3 study of cosmic homogeneity using BOSS DR12 quasar sample*, **11**, 060 (2016), doi:10.1088/1475-7516/2016/11/060, arXiv:1602.09010.

[43] S. M. Croom, B. J. Boyle, T. Shanks, R. J. Smith, *et al.*, *The 2dF QSO Redshift Survey - XIV. Structure and evolution from the two-point correlation function*, **356**, 415 (2005), doi:10.1111/j.1365-2966.2004.08379.x, astro-ph/0409314.

[44] N. Kaiser, *Clustering in real space and in redshift space*, **227**, 1 (1987).

[45] É. Aubourg, S. Bailey, J. E. Bautista, F. Beutler, *et al.* (????).

[46] A. Lewis et S. Bridle, *Cosmological parameters from CMB and other data : A Monte Carlo approach*, **66**, 103511 (2002), doi:10.1103/PhysRevD.66.103511, astro-ph/0205436.

[47] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, *et al.*, *Planck 2015 results. XIV. Dark energy and modified gravity*, **594**, A14 (2016), doi:10.1051/0004-6361/201525814, arXiv:1502.01590.

[48] W. Hu et M. White, *Acoustic Signatures in the Cosmic Microwave Background*, **471**, 30 (1996), doi:10.1086/177951, astro-ph/9602019.

[49] M. Kunz, P.-S. Corasaniti, D. Parkinson, et E. J. Copeland, *Model-independent dark energy test with σ_8 using results from the Wilkinson Microwave Anisotropy Probe*, **70**, 041301 (2004), doi:10.1103/PhysRevD.70.041301, astro-ph/0307346.

[50] M. Baldi et V. Pettorino, *High-z massive clusters as a test for dynamical coupled dark energy*, **412**, L1 (2011), doi:10.1111/j.1745-3933.2010.00975.x, arXiv:1006.3761.

[51] V. Acquaviva et C. Baccigalupi, *Dark energy records in lensed cosmic microwave background*, **74**, 103510 (2006), doi:10.1103/PhysRevD.74.103510, astro-ph/0507644.

[52] C. Carbone, M. Baldi, V. Pettorino, et C. Baccigalupi, *Maps of CMB lensing deflection from N-body simulations in Coupled Dark Energy Cosmologies*, **9**, 004 (2013), doi:10.1088/1475-7516/2013/09/004, arXiv:1305.0829.

[53] L. A. Kofman et A. A. Starobinskii, *Effect of the Cosmological Constant on Largescale Anisotropies in the Microwave Background*, Soviet Astronomy Letters **11**, 271 (1985).

[54] T. Giannantonio, R. Scranton, R. G. Crittenden, R. C. Nichol, *et al.*, *Combined analysis of the integrated Sachs-Wolfe effect and cosmological implications*, **77**, 123520 (2008), doi:10.1103/PhysRevD.77.123520, arXiv:0801.4380.

[55] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, *et al.*, *Binary Black Hole Mergers in the First Advanced LIGO Observing Run*, Physical Review X **6**, 041015 (2016), doi:10.1103/PhysRevX.6.041015, arXiv:1606.04856 [gr-qc].

[56] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, *et al.*, *GW170817 : Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, Physical Review Letters **119**, 161101 (2017), doi:10.1103/PhysRevLett.119.161101, arXiv:1710.05832 [gr-qc].

[57] A. Goldstein, P. Veres, E. Burns, M. S. Briggs, *et al.*, *An Ordinary Short Gamma-Ray Burst with Extraordinary Implications : Fermi-GBM Detection of GRB 170817A*, **848**, L14 (2017), doi:10.3847/2041-8213/aa8f41, arXiv:1710.05446 [astro-ph.HE].

[58] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, *et al.*, *Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger : GW170817 and GRB 170817A*, **848**, L13 (2017), doi:10.3847/2041-8213/aa920c, arXiv:1710.05834 [astro-ph.HE].

[59] P. Creminelli et F. Vernizzi, *Dark energy after gw170817 and grb170817a*, Phys. Rev. Lett. **119**, 251302 (2017), doi:10.1103/PhysRevLett.119.251302.

[60] J. M. Ezquiaga et M. Zumalacárregui, *Dark energy after gw170817 : Dead ends and the road ahead*, Phys. Rev. Lett. **119**, 251304 (2017), doi:10.1103/PhysRevLett.119.251304.

[61] T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, *et al.*, *Strong constraints on cosmological gravity from gw170817 and grb 170817a*, Phys. Rev. Lett. **119**, 251301 (2017), doi:10.1103/PhysRevLett.119.251301.

[62] J. Sakstein et B. Jain, *Implications of the neutron star merger gw170817 for cosmological scalar-tensor theories*, Phys. Rev. Lett. **119**, 251303 (2017), doi:10.1103/PhysRevLett.119.251303.

[63] E.-M. Mueller, W. Percival, E. Linder, S. Alam, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : constraining modified gravity*, **475**, 2122 (2018), doi:10.1093/mnras/stx3232, arXiv:1612.00812.

[64] E. V. Linder et R. N. Cahn, *Parameterized beyond-Einstein growth*, Astroparticle Physics **28**, 481 (2007), doi:10.1016/j.astropartphys.2007.09.003, astro-ph/0701317.

[65] Y. Gong, *Growth factor parametrization and modified gravity*, **78**, 123010 (2008), doi:10.1103/PhysRevD.78.123010, arXiv:0808.1316.

[66] J. Li et G.-B. Zhao, *Cosmological tests of gravity with latest observations*, ArXiv e-prints (2018), arXiv:1806.05022.

[67] J. N. Grieb, A. G. Sánchez, S. Salazar-Albornoz, R. Scoccimarro, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : Cosmological implications of the Fourier space wedges of the final sample*, **467**, 2085 (2017), doi:10.1093/mnras/stw3384, arXiv:1607.03143.

[68] A. G. Sánchez, R. Scoccimarro, M. Crocce, J. N. Grieb, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : Cosmological implications of the configuration-space clustering wedges*, **464**, 1640 (2017), doi:10.1093/mnras/stw2443, arXiv:1607.03147.

[69] J. Neveu, V. Ruhlmann-Kleider, P. Astier, M. Besançon, *et al.*, *Constraining the ΛCDM and Galileon models with recent cosmological data*, **600**, A40 (2017), doi:10.1051/0004-6361/201628878, arXiv:1605.02627 [gr-qc].

[70] A. Vainshtein, *To the problem of nonvanishing gravitation mass*, Physics Letters B **39**, 3 (1972).

[71] A. Barreira, B. Li, W. A. Hellwing, C. M. Baugh, *et al.*, *Nonlinear structure formation in the cubic Galileon gravity model*, Journal of Cosmology and Astroparticle Physics **2013**, 027 (2013), doi:10.1088/1475-7516/2013/10/027, arXiv:1306.3219.

[72] B. Li, A. Barreira, C. M. Baugh, W. A. Hellwing, *et al.*, *Simulating the quartic Galileon gravity model on adaptively refined meshes*, Journal of Cosmology and Astroparticle Physics **2013**, 012 (2013), doi:10.1088/1475-7516/2013/11/012, 1308.3491.

[73] A. Barreira, B. Li, C. M. Baugh, et S. Pascoli, *Spherical collapse in Galileon gravity : fifth force solutions, halo mass function and halo bias*, Journal of Cosmology and Astroparticle Physics **2013**, 056 (2013), doi:10.1088/1475-7516/2013/11/056, arXiv:1308.3699.

[74] G.-B. Zhao, T. Giannantonio, L. Pogosian, A. Silvestri, *et al.*, *Probing modifications of general relativity using current cosmological observations*, **81**, 103510 (2010), doi:10.1103/PhysRevD.81.103510, arXiv:1003.0001 [astro-ph.CO].

[75] E. Bertschinger et P. Zukin, *Distinguishing modified gravity from dark energy*, **78**, 024015 (2008), doi:10.1103/PhysRevD.78.024015, arXiv:0801.2431.

[76] T. Giannantonio, M. Martinelli, A. Silvestri, et A. Melchiorri, *New constraints on parametrised modified gravity from correlations of the CMB with large scale structure*, **4**, 030 (2010), doi:10.1088/1475-7516/2010/04/030, arXiv:0909.2045.

[77] Y. Wang, G.-B. Zhao, C.-H. Chuang, M. Pellejero-Ibanez, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : a tomographic analysis of structure growth and expansion rate from anisotropic galaxy clustering*, ArXiv e-prints (2017), arXiv:1709.05173.

[78] G.-B. Zhao, *Modeling the Nonlinear Clustering in Modified Gravity Models. I. A Fitting Formula for the Matter Power Spectrum of f(R) Gravity*, **211**, 23 (2014), doi:10.1088/0067-0049/211/2/23, arXiv:1312.1291.

Chapitre 6

Conclusion and prospects

Contents

One of the biggest questions of contemporary cosmology is the origin of late-time cosmic acceleration and the formation of large-scale structures. Chapter 1 presents the ingredients of the standard cosmological model called ΛCDM and the different surveys whose major goal is to probe the nature of the cosmic acceleration. One approach consists in measuring cosmological distances using the baryon acoustic oscillations (BAO) as a standard ruler. This technique has been shown to be very robust towards non-linearities and redshift space distortions (RSD). A complementary approach is to measure the growth of structure by modeling the full shape of the two-point correlation function using the RSD technique. The growth rate can be used as a direct test of the underlying theory of gravitation, general relativity (GR). Any departure from the ΛCDM+GR prediction would imply modifications of the theory at cosmological scales.

6.1 Conclusions of this work

Among the ongoing surveys, eBOSS is the fourth iteration of SDSS and started taking data in 2014. It extends the BOSS spectroscopic survey with two new tracers : quasars in the intermediate redshift $0.8 \leq z \leq 2.2$ and emission line galaxies (ELG) in the redshift range $0.6 < z < 1.1$. Contrary to BOSS where quasars were observed at high redshift $z > 2.1$ to use the Lyman-α forests in quasar spectra as tracers of the neutral hydrogen in the intergalactic medium, eBOSS uses quasars themselves at lower redshifts as direct tracers of the matter field.

During my thesis work, I worked on a key science goal of eBOSS, namely the analysis of the anisotropic clustering in configuration space of the eBOSS DR14 quasar sample. This sample represents two years of data from eBOSS and includes 148,659 quasars spread over the redshift range $0.8 \leq z \leq 2.2$ and spanning 2112.9 square degrees. In chapter 2, we present the SDSS-IV eBOSS experiment, its survey strategy including the target selec-

tion and the spectroscopic pipeline to obtain a catalogue of all spectroscopically-observed quasars in the redshift range of interest. We also investigate the issue of redshift uncertainties, as emission lines in quasar spectra are shifted due to matter outflows around the super-massive black hole. It is therefore necessary to control the impact of redshift uncertainties on the cosmological parameters. Then, we explain the construction of the large-scale structure catalogues for clustering analysis which takes into account the missing targets by resorting to a weighting scheme.

Chapter 3 reviews the theoretical formalism used to predict the two-point correlation function in redshift space. Three types of non-linearities at least are challenging to model : the non-linear evolution of density and velocity fields, the non-linear relation between dark matter and quasars and the non-linear mapping from real to redshift space. As a consequence, the RSD technique requires more efforts to evaluate the performance of the modeling, compared to the BAO-only method.

In chapter 4, we describe the methodology with the RSD model, the decomposition of the two-point correlation function into the first three even Legendre multipoles and into three wedges, the covariance matrix and the way parameters are inferred. Our RSD model uses the Convolution Lagrangian Perturbation Theory (CLPT) with a Gaussian Streaming (GS) model and we demonstrate its applicability for dark matter halos of masses of the order of $10^{12.5}\mathrm{M}_{\odot}$ hosting eBOSS quasar tracers at mean redshift $z \simeq 1.5$. In order to estimate the systematic error budget related to the RSD modeling, we use the N-body Outer Rim simulation to test the predictions of CLPT in real space and then evaluate the performance of the model in redshift space using a hundred mock catalogues created for that purpose. We investigate both the effect of the bias model and the spectroscopic redshift resolution in the RSD modeling. The eBOSS quasar sample suffers from an important systematic uncertainty related to spectroscopic redshift precision : we study its effect by modeling a Gaussian redshift resolution and a more physical resolution using the comparison between different redshift estimates, z and $\mathrm{z_{MgII}}$. We demonstrate that accounting for the non-Gaussian tails of the physical distributions has a sizeable impact on the response of the model. In fact, about half of the quoted uncertainty on $\Delta f\sigma_8$ arises from redshift resolution effects. In this analysis, we propose a way of investigating spectroscopic redshift resolution using mock catalogues. We also propose to move beyond the traditional weighting scheme that was used for BOSS galaxies and the BAO measurement with the DR14 sample to account for redshift failures and close-pairs. We validate the procedure on a thousand of approximate EZ mock catalogues. This approach allows the observational systematics to be much smaller than the current statistical precision, this is why eventually we only report a systematic uncertainty related to the modeling. We also perform a series of consistency checks on the data, including tests on the weighting scheme, change in the covariance matrix and in the fitting conditions.

Chapter 5 presents the main results of this thesis work. The full-shape (FS) analysis in configuration space and BAO-only analysis have been published in [1, 2].
Regarding the FS analysis, we first check that the multipoles and wedges approaches yield consistent results. The decomposition into Legendre multipoles provides the cosmological measurements with the best statistical precision. At the effective redshift $z_{\mathrm{eff}} = 1.52$, the growth rate of structures $f\sigma_8(z_{\mathrm{eff}}) = 0.426 \pm 0.077$, the expansion rate $H(z_{\mathrm{eff}}) = 159^{+12}_{-13}(r_s^{\mathrm{fid}}/r_s)\mathrm{km.s^{-1}.Mpc^{-1}}$, and the angular diameter distance $D_A(z_{\mathrm{eff}}) = 1850^{+90}_{-115}\,(r_s/r_s^{\mathrm{fid}})\mathrm{Mpc}$ where r_s is the sound horizon at the end of the baryon drag epoch and r_s^{fid} is its value in the fiducial cosmology. The quoted uncertainties include both systematic and statistical contributions. We also split the redshift sample into two bins and find consistent results with a redshift evolution for the bias and the redshift uncertainties, but the cur-

rent statistics is not enough to be sensitive to the redshift evolution of the cosmological parameters. We find consistent results with our companion paper in Fourier space [3] both at effective redshift $z_{\rm eff} = 1.52$ and for the two bins in redshift. New techniques that account for the redshift evolution applying optimal redshift weights [4, 5] are also described and a comparison between the five RSD analysis [1, 3, 6, 4, 5] is proposed. The results presented in this work are found to be in agreement with the other companion papers using the same data sample but analysed with different techniques, demonstrating the complementary and the robustness of each method. Regarding the BAO-only analysis, it corresponds to the first detection of the BAO in a quasar sample at intermediate redshifts $1 < z < 2$. We obtain a 3.8% measurement on the spherically-averaged BAO distance $D_V(z_{\rm eff} = 1.52) = 3843 \pm 147(r_d/r_s^{\rm fid})$ and it is consistent with the measurement of D_V from FS analysis. Using the BAO data alone from our work and previous independent BAO measurements from BOSS galaxies [7] and Ly-α forests [8, 9], we tested a ΛCDM model with free curvature, assuming only that the acoustic scale has a fixed comoving size. We find $\Omega_\Lambda > 0$ at 6.6σ significance. The results on the evolution of distances from BAO and are consistent with the predictions of ΛCDM with *Planck* parameters assuming the existence of a cosmological constant to explain the late-time acceleration of the expansion of the Universe. The measurement of $f\sigma_8$ is consistent with General Relativity (GR) in the almost unexplored redshift range probed by the eBOSS quasar sample. Previous works have started exploring the $z > 0.8$ range using ELG [FastSound, 10] and from a multi-sample of galaxies [VIPERS, 11] and obtained a $\sim$25% measurement of $f\sigma_8$ using the monopole and the quadrupole at fixed H and D_A. In this thesis work, we measure simultaneously $f\sigma_8$, D_A and H and obtain a 18% measurement of $f\sigma_8$ after marginalizing over the full set of parameters. When fixing D_A and H, we measure $f\sigma_8$ with 11% precision. Therefore, this work improves the precision of the cosmological parameters, but also extends the inferred cosmological parameters and provides a measurement of the growth rate of structure that can be used to extend the tests of modified gravity models at higher redshift ($z > 1$). We emphasize that measurements of $f\sigma_8$ at fixed D_A and H obtain smaller uncertainties that do not account for the marginalization over the full set of parameters and hence cannot be used to test alternative scenarios of gravity in general. In chapter 5, we also discuss the cosmological implications of the eBOSS DR14 quasar sample. The combination of BOSS DR12 galaxies, BOSS DR12 Ly-α, eBOSS DR14 quasars [1] and *Planck* [12] yields the following constraints : $\Omega_m = 0.3094^{+0.0076}_{-0.0080}$, $\Omega_\Lambda = 0.697^{+0.035}_{-0.032}$ and hence the curvature density $\Omega_k = -0.007 \pm 0.030$. The results are in agreement with a universe dominated by dark energy with no curvature. Moreover, we test the ΛCDM+GR prediction for $f\sigma_8$ where $f = \Omega_m^\gamma$ with $\gamma = 0.55$ by allowing γ to vary. Our measurement alone is not sensitive enough to provide competitive constraint on γ. When combining *Planck* with BOSS DR12 galaxies and eBOSS DR14 quasars, we find $\gamma = 0.55 \pm 0.19$, and when combining with additional probes such as Type Ia supernovae, weak lensing and measurements of H_0, we find $\gamma = 0.485 \pm 0.031$ [13]. We also provide a discussion on the cosmological implications in the framework of alternative gravity models for the Galileon and scalar-tensor theories, including a f(R) theory that mimics the background evolution of ΛCDM. This specific f(R) is found to be strongly disfavored by current data and might be even ruled out with the constraints from the final eBOSS samples.

6.2 Prospects

6.2.1 eBOSS prospects : analysis of the final quasar sample

This study is a first use of eBOSS quasars for BAO-only and full-shape analyses and will be included for the final eBOSS sample. The final projected area for the full quasar sample will correspond to 5500 square degrees compared to the 2213 square degrees of the DR14 sample. eBOSS observations will end in February 2019 and we are planning to release the final catalogue, DR16, along with the BAO and full-shape analyses by the end of 2019. The following table summarizes the expected statistical precision on D_V, D_A, H and $f\sigma_8$, which can be obtained by re-scaling the current precision given the projected final area :

Sample	area (deg^2)	σ_{D_V}/D_V	σ_{D_A}/D_A	σ_H/H	$\sigma_{f\sigma_8}/f\sigma_8$
DR14	2213	3.8%	5.5%	7.9%	16.4%
DR16	5500	2.4%	3.4%	4.9%	10.1%

Given the statistical precision of the current DR14 quasar sample, the reported systematic uncertainty is not dominant in our analysis, but further investigations including a full blind mock challenge similar to that undertaken for BOSS is in progress and will be available in time for the analysis of the final eBOSS sample. In parallel, further improvements in the model to take into account the shape of the redshift uncertainties distribution are also considered to reduce the systematic error budget. In fact, if we allow the systematic budget to represent 20% of the statistical precision of the DR16 sample,the requirement on the allowed shift for $f\sigma_8$ is $\Delta f\sigma_8 < 0.01$. It implies a reduction of a factor 2-3 of the current systematic uncertainty related to the modeling. There is therefore an important effort within the galaxy-quasar working group to prepare a mock challenge to test modeling and observational effects using different RSD models. Among the publicly available models for RSD, the following will be considered :

— In Fourier space : an extension of the TNS model [14] using 2-loop resummed perturbation theory (RPT) as in our companion paper [3], or using regularized perturbation theory [15, RegPT,]

— In configuration space : the CLPT-GS model used in this work [1], the CLEFT-GS model [16] which is an extension of CLPT including additional bias terms and effective contributions. We can also test the TNS model in configuration space in association with RegPT or with gRPT (Crocce, Blas and Scoccimarro in prep.) as in [6].

Regarding redshift quasar uncertainties, there are ongoing efforts within the pipeline team to provide new pipeline redshifts. Their performance in terms of catastrophic redshifts rate and resolution will be evaluated and the visual inspection procedure will be determined accordingly. Eventually, the direct use of quasars as tracers is not the only study using eBOSS data. The BAO-only analysis of the luminous red galaxies (LRG) at an effective redshift $z_{\rm eff} = 0.72$ yields a 2.6% measurement on D_V and the RSD analysis will be published soon. Ongoing analyses are focused on ELG at $z \sim 0.8$ and preliminary results on mock catalogues without systematic effects find H, D_A and $f\sigma_8$ with respectively 6.1%, 4.6% and 14% precision. The improvement of statistics would also allow different methods and tracers to be combined and thus to provide even tighter constraints on cosmological parameters.

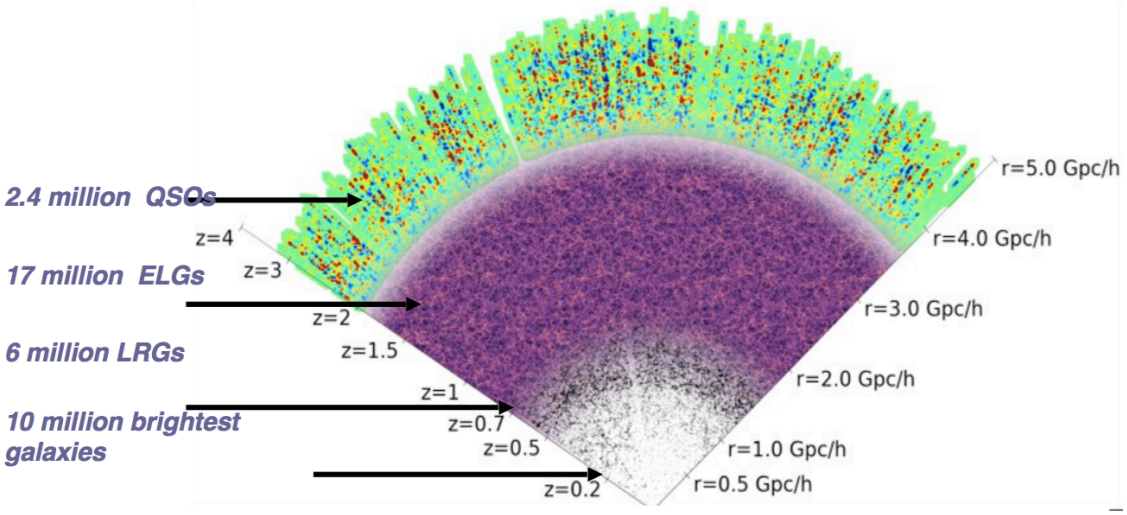


FIGURE 6.1 – The five cosmological tracers of DESI that will cover the redsfhit range up to $z = 3.5$.

6.2.2 Prospects for stage IV experiments

Together with BOSS, eBOSS is of particular interest since it paves the way for future programs such as the ground-based Dark Energy Spectroscopic Instrument [DESI, 17, 18], the space-based mission, *Euclid* [19] and the Large Synoptic Survey Telescope [LSST, 20]. These Stage-IV experiments will mark the beginning of an era of precision with sub-percent measurements of cosmological parameters across the redshift.

— DESI is a spectroscopic survey (2020-2025) that will cover 14,000 deg^2 and that will use five tracers as illustrated in figure 6.1 : a new sample of ∼10 million bright galaxies (BGS) at low redshift $z < 0.5$, 6 million LRG between $0.4 < z < 1.0$, 17 million ELG between $0.6 < z < 1.6$, 1.7 million quasars between $0.9 < z < 2.1$ and 0.7 million Ly-α forests at $z > 2.1$.

DESI will therefore extensively probe the intermediate redshift range $1 < z < 2$ with millions of spectra, pushing an order of magnitude beyond current measurements from BAO and RSD techniques. Its science goals is to measure the distance scale from BAO to sub-percent levels all across the redshift range. The top panel of figure 6.2shows the DESI forecasts for the expansion rate $H(z)/(1 + z)$ as a function of redshift and the bottom panel represents the DESI forecasts for the growth rate as a function of redshift. We add the forecasts for the final eBOSS samples in blue for comparison. We stress that the reported precision for eBOSS is obtained at the effective redshift of each sample, while DESI will have a better precision across narrow redshift bins of width $\Delta z = 0.1$.

— The *Euclid* mission (2021-2026) has been optimized for two complementary cosmological probes : BAO and weak gravitational lensing. It will produce a catalogue of up to 100 million galaxy redshifts and an imaging survey that will estimate the galaxy ellipticity of up to 2 billion galaxy images. The survey will cover ∼15,000 deg^2 and will target Hα emitters since they guarantee both relatively dense sampling and an efficient method to measure redshifts. In particular, *Euclid* will observe 50 million ELG in the redshift range $0.8 < z < 1.8$. The *Euclid* scientific goals is to measure the dark energy equation of state $w = w_0 + (1 - a)w_a$ with 2% precision on w_0 and 10% on w_a, improving the current constraints by a factor of 10 and 50 at least.

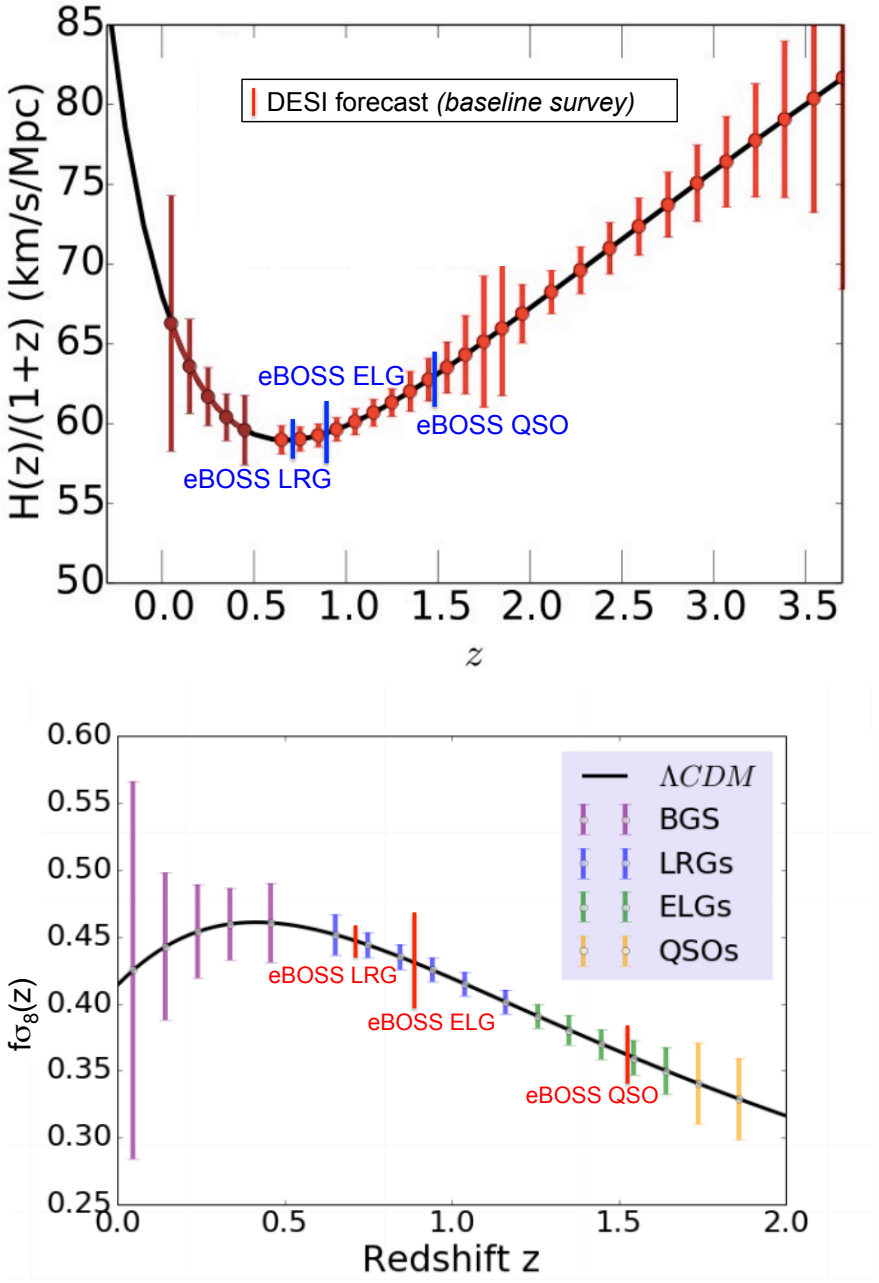


Figure 6.2 – Top panel : The DESI forecasts (in red) for the baseline survey, i.e. 14,000 deg^2, for the expansion rate as a function of redsfhit. The eBOSS DR16 forecasts from LRG at $z_{\rm eff} = 0.72$, ELG at $z_{\rm eff} = 0.86$ and quasars at $z_{\rm eff} = 1.52$ are plotted in blue. We emphasize that the eBOSS forecasts are obtained by averaging over the redshift range while DESI will measure sub-percents across redshift bins of width $\Delta z = 0.1$. Bottom panel : Same for $f\sigma_8$.

The *Euclid* scientific goals also include a measurement of the growth index γ with a precision of 2%, improving the current constraint by a factor 30. In the context of dark energy and as a test of general relativity in particular, it would also be highly valuable to combine weak gravitational lensing and galaxy redshift-space clustering to break degeneracies between the growth rate of structure and the amplitude of dark matter fluctuations and to measure the gravitational slip parameter E_G defined as the ratio between the two gravitational potentials Ψ and Φ. In the standard cosmological model, this parameter should asymptote towards a given value on large linear scales. If such a test fails, it would either imply an incorrect matter energy density or a departure from standard gravity.

— LSST (2022-2032) is an optical/near-infrared imaging survey of half the sky (20,000 deg^2 in the South) in *ugrizy* bands to $r \sim 27.5$ (100 times fainter than SDSS) which will produce a catalogue of 20 billion stars and 20 billion galaxies. After ten years, half of the sky will be imaged about 1000 times, providing exquisite photometry, astrometry and image quality. LSST and Euclid will have an overlapping area of ∼10,000 deg^2 which will allow us to combine measurements and reduce the impact of potential systematics.

On one hand, the precision of DESI and *Euclid* will be sufficient to highlight a possible time-evolution of dark energy and to discriminate between dark energy and modified-gravity theories. Both will target with the highest priority the intermediate redshift range $1 < z < 2$ as it corresponds to a transition to the present epoch dominated by cosmic acceleration. On the other hand, *Euclid* and LSST will extensively map the matter distribution using the weak lensing technique. We mentioned above the example of the E_G parameter as a promising observable that combines both galaxy clustering and weak lensing statistics. There is no doubt that additional tests of ΛCDM will emerge as a result of the combination of independent probes which will reduce systematics and break degeneracies in the cosmological information we can extract. The bottom line is that we are really entering an area of precision cosmology with a huge amount of data coming from complementary facilities and I am very enthousiastic about developing new skills that will allow me contributing to the synergy among the future surveys.

Bibliographie

[1] P. Zarrouk, E. Burtin, H. Gil-Marín, A. J. Ross, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2*, **477**, 1639 (2018), doi:10.1093/mnras/sty506.

[2] M. Ata, F. Baumgarten, J. Bautista, F. Beutler, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : First measurement of Baryon Acoustic Oscillations between redshift 0.8 and 2.2*, ArXiv e-prints (2017), arXiv:1705.06373.

[3] H. Gil-Marín, J. Guy, P. Zarrouk, E. Burtin, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : structure growth rate measurement from the anisotropic quasar power spectrum in the redshift range 0.8 z 2.2*, **477**, 1604 (2018), doi:10.1093/mnras/sty453, arXiv:1801.02689.

[4] R. Ruggeri, W. J. Percival, H. Gil Marin, F. Beutler, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : measuring the evolution of the growth rate using redshift space distortions between redshift 0.8 and 2.2*, ArXiv e-prints (2018), arXiv:1801.02891.

[5] G.-B. Zhao, Y. Wang, S. Saito, H. Gil-Marín, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : a tomographic measurement of cosmic structure growth and expansion rate based on optimal redshift weights*, ArXiv e-prints (2018), arXiv:1801.03043.

[6] J. Hou, A. G. Sánchez, R. Scoccimarro, S. Salazar-Albornoz, *et al.*, *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : anisotropic clustering analysis in configuration-space*, ArXiv e-prints (2018), arXiv:1801.02656.

[7] S. Alam, M. Ata, S. Bailey, F. Beutler, *et al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey : cosmological analysis of the DR12 galaxy sample*, **470**, 2617 (2017), doi:10.1093/mnras/stx721, arXiv:1607.03155.

[8] J. E. Bautista, N. G. Busca, J. Guy, J. Rich, *et al.*, *Measurement of baryon acoustic oscillation correlations at $z = 2.3$ with SDSS DR12 Lyα-Forests*, **603**, A12 (2017), doi:10.1051/0004-6361/201730533, arXiv:1702.00176.

[9] H. du Mas des Bourboux, J.-M. Le Goff, M. Blomqvist, N. G. Busca, *et al.*, *Baryon acoustic oscillations from the complete SDSS-III Lyα-quasar cross-correlation function at $z = 2.4$*, ArXiv e-prints (2017), arXiv:1708.02225.

[10] T. Okumura, C. Hikage, T. Totani, M. Tonegawa, *et al.*, *The Subaru FMOS galaxy redshift survey (FastSound). IV. New constraint on gravity theory from redshift space distortions at z 1.4*, **68**, 38 (2016), doi:10.1093/pasj/psw029, arXiv:1511.08083.

[11] A. Pezzotta, S. de la Torre, J. Bel, B. R. Granett, *et al.*, *The VIMOS Public Extragalactic Redshift Survey (VIPERS). The growth of structure at 0.5 < z < 1.2 from redshift-space distortions in the clustering of the PDR-2 final sample*, **604**, A33 (2017), doi:10.1051/0004-6361/201630295.

[12] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, *et al.*, *Planck 2015 results. XIII. Cosmological parameters*, **594**, A13 (2016), doi:10.1051/0004-6361/201525830, arXiv:1502.01589.

[13] J. Li et G.-B. Zhao, *Cosmological tests of gravity with latest observations*, ArXiv e-prints (2018), arXiv:1806.05022.

[14] A. Taruya, T. Nishimichi, et S. Saito, *Baryon acoustic oscillations in 2D : Modeling redshift-space power spectrum from perturbation theory*, **82**, 063522 (2010), doi:10.1103/PhysRevD.82.063522, arXiv:1006.0699.

[15] A. Taruya, F. Bernardeau, T. Nishimichi, et S. Codis, *Direct and fast calculation of regularized cosmological power spectrum at two-loop order*, **86**, 103528 (2012), doi:10.1103/PhysRevD.86.103528, arXiv:1208.1191 [astro-ph.CO].

[16] Z. Vlah, E. Castorina, et M. White, *The Gaussian streaming model and convolution Lagrangian effective field theory*, **12**, 007 (2016), doi:10.1088/1475-7516/2016/12/007, arXiv:1609.02908.

[17] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, *et al.*, *The DESI Experiment Part I : Science,Targeting, and Survey Design*, ArXiv e-prints (2016), arXiv:1611.00036 [astro-ph.IM].

[18] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, *et al.*, *The DESI Experiment Part II : Instrument Design*, ArXiv e-prints (2016), arXiv:1611.00037 [astro-ph.IM].

[19] L. Amendola, S. Appleby, D. Bacon, T. Baker, *et al.*, *Cosmology and Fundamental Physics with the Euclid Satellite*, Living Reviews in Relativity **16**, 6 (2013), doi:10.12942/lrr-2013-6, arXiv:1206.1225.

[20] LSST Dark Energy Science Collaboration, *Large Synoptic Survey Telescope : Dark Energy Science Collaboration*, ArXiv e-prints (2012), arXiv:1211.0310 [astro-ph.CO].

Résumé substantiel

6.3 Introduction

La cosmologie a pour but de retracer l'histoire de l'univers en s'appuyant à la fois sur des modèles mathématiques et sur des observations. Dès les années 1920, l'analyse du mouvement des galaxies, ces regroupements d'étoiles attirées sous l'effet de leur masse, ont montré que les raies spectrales de leurs spectres se décalaient vers le rouge. Cet effet, appelé redshift [1] (ou décalage vers le rouge en français) est d'origine cosmologique : les galaxies s'éloignent les unes des autres car l'univers (i.e. l'espace lui-même) est en expansion. A la fin des années 90, d'autres observations ont révélé que les distances entre les galaxies augmentent plus rapidement les derniers 6 milliards d'années de l'univers, lorsque nous estimons que notre univers est âgé de 13,8 milliards d'années. Cette découverte capitale de l'accélération de l'expansion de l'univers est considérée comme l'une des plus grandes énigmes de la cosmologie et est au cœur de mon sujet de thèse.

6.3.1 Le modèle ΛCDM

Les observations du fond diffus cosmologique (CMB, Cosmic Microwave Background) ont révélé que l'univers primordial était beaucoup plus chaud, plus dense et plus uniforme que l'univers actuel, avec des fluctuations de température de seulement 10^{-5} K. Ces fluctuations primordiales de température, et donc de densité, ont ensuite grandi par effondrement gravitationnel dans un univers en expansion pour former une hiérarchie de structures liées par interaction gravitationnelle jusqu'à représenter une toile cosmique que nous observons aujourd'hui. Cette histoire de l'univers est retracée sur la droite de la figure 6.3 depuis un univers primordial dominé par le rayonnement puis la matière avec formation de structures jusqu'à la phase d'expansion accélérée qui dure depuis 6 milliards d'années.

Afin d'expliquer ces observations, le modèle actuel de la cosmologie, appelé modèle ΛCDM, repose sur les éléments suivants :

- l'univers est homogène et isotrope à très grande échelle
- la gravitation est décrite par la théorie de la relativité générale
- le modèle dépend de 6 paramètres : la densité baryonique Ω_b, la densité de matière noire Ω_{cdm}, la densité d'énergie noire Ω_Λ, l'indice spectral des perturbations primordiales scalaires n_s, l'amplitude des perturbations primordiales de courbure A_s et l'épaisseur optique de réionisation τ.

Le contenu énergétique de l'univers aujourd'hui est représenté à gauche de la figure 6.3 avec 5% de matière ordinaire (étoiles, galaxies, gaz, poussière, nous, ...) ; 25% de

1. Le redshift cosmologique est défini par : $1+z = \frac{\lambda_{\text{obs}}}{\lambda_{\text{R.F.}}} = \frac{1}{a(t)}$ où λ_{obs} est la longueur d'onde observée, $\lambda_{\text{R.F.}}$ est la longueur d'onde de référence au repos et $a(t)$ est le facteur d'échelle qui quantifie le taux d'expansion de l'univers.

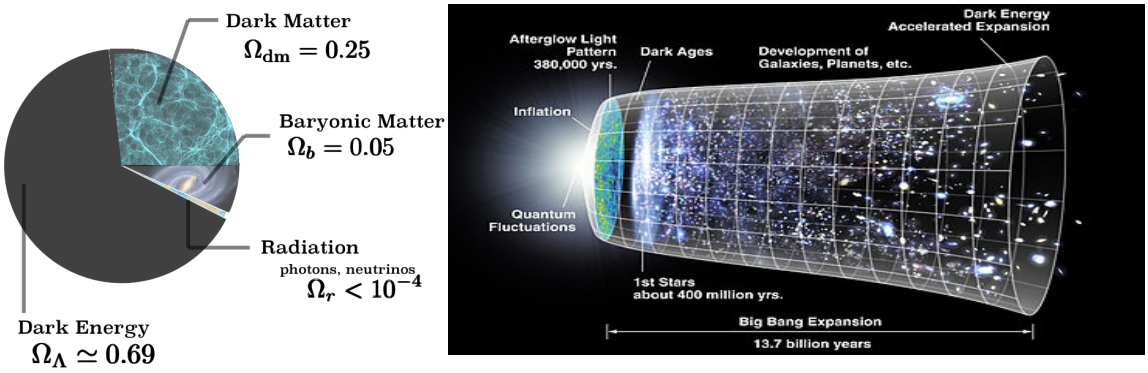


FIGURE 6.3 – Gauche : contenu énergétique actuel de l'univers. Droite : Frise retraçant l'histoire de l'univers depuis le Big Bang il y a 13.7 milliards d'années jusqu'à aujourd'hui.

matière noire autrement dit des particules massives froid (Cold Dark Matter, CDM) non sensibles à l'interaction électromagnétique et qui doivent interagir très peu avec la matière ordinaire pour rendre compte des observations de la formation des structures ; et enfin 70% d'énergie noire un fluide exotique aux propriétés constantes dans le temps et l'espace, aussi appelé constante cosmologique et qui serait responsable de l'accélération de l'expansion de l'univers.

Dans le cadre de ma thèse, je m'intéresse à la question de la nature de l'énergie noire qui domine le contenu énergétique de l'univers aujourd'hui. Une autre alternative à l'énergie noire pour expliquer la phase d'expansion accélérée de l'univers serait de modifier la théorie de la relativité générale aux échelles cosmologiques, là où nous supposons qu'elle est valable alors qu'elle est essentiellement testée, et ce avec une précision remarquable, aux échelles du système solaire. A ce jour, aucune théorie ne permet d'expliquer de manière satisfaisante l'origine de l'accélération de l'expansion de l'univers, d'où le nombre croissant de projets qui y sont consacrés.

6.3.2 Cartographier les structures de l'univers

Dans le cadre de ma thèse, je fais partie de l'un des plus ambitieux programmes d'observation, le programme SDSS (Sloan Digital Sky Survey) qui utilise un télescope de 2,5 mètres de diamètre, situé à l'observatoire de Apache Point au Nouveau Mexique, États-Unis. Le télescope est en opération depuis les années 2000 et a déjà permis de collecter la position dans le ciel de plus d'un million de galaxies.

La figure de gauche 6.4 montre une carte des structures obtenue avec le SDSS où l'observateur est au centre et chaque point correspond à une galaxie repérée par ses deux coordonnées angulaires (ascension droite RA et déclinaison DEC) et par le redshift z. On observe que les galaxies ne sont pas distribuées de manière uniforme, certaines régions sont plus denses, d'autres semblent vide de matière. Pour décrire la distribution des structures, nous définissons un champ de densité de contraste $\delta(x)$ défini par :

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}} \tag{6.1}$$

où $\bar{\rho}$ correspond à la densité moyenne de l'univers. La distribution des corrélations spatiales des galaxies peut alors être étudiée au moyen d'outils statistiques, en particulier la fonction de corrélation à deux points $xi(r)$ définie par :

$$\xi(r) = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) \rangle \tag{6.2}$$

Cet outil statistique quantifie le degré de structuration de la matière, une autre interprétation est de considérer la fonction de corrélation comme l'excès de probabilité par rapport à une distribution uniforme de trouver deux galaxies séparées par une distance r.

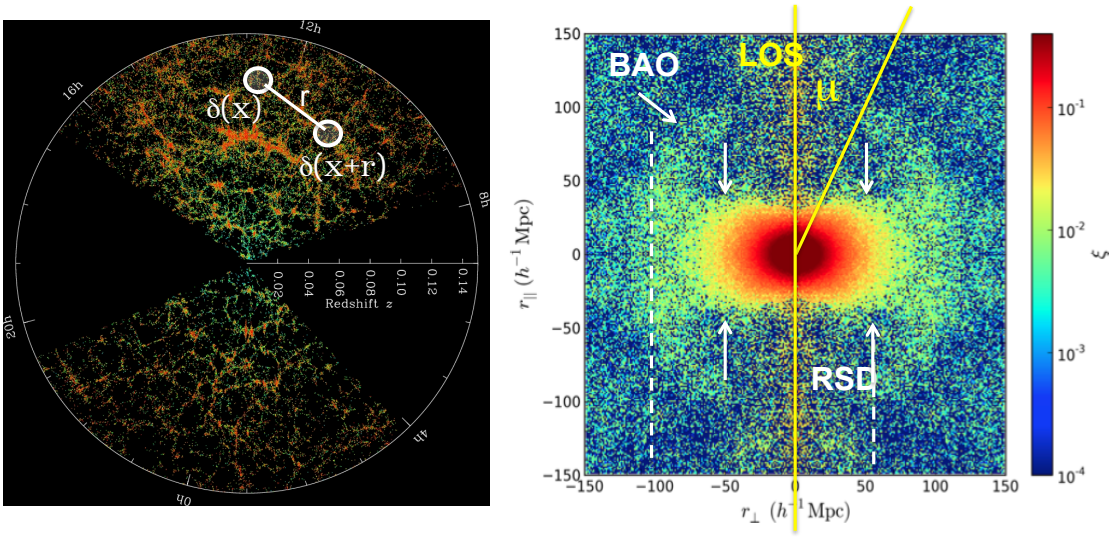


FIGURE 6.4 – Gauche : carte des structures obtenue avec le SDSS où l'observateur est au centre et chaque point représente une galaxie. Droite : Fonction de corrélation à deux points représentée dans les directions parallèle et perpendiculaire à la ligne de visée.

La mesure de cette distribution au moyen de grands relevés comme le SDSS a ouvert une nouvelle piste pour explorer la nature de l'énergie noire grâce à l'exploitation d'une nouvelle sonde : les oscillations acoustiques de baryons (BAO pour Baryon Acoustic Oscillations). Il s'agit d'ondes d'ondes de pression formées par le plasma primordial de baryons et de photons lorsque que l'univers était plus chaud et dense que l'univers que nous observons aujourd'hui. Ces ondes sonores qui se sont propagées pendant les 380 000 premières années de l'univers, se sont figées au moment où les baryons se sont couplés aux électrons pour former des atomes d'hydrogène et ont donc cessé d'interagir avec les photons dans le plasma primordial. L'arrêt de ces oscillations a laissé une empreinte sous forme d'une distance caractéristique dans la distribution actuelle de la matière, appelée « pic BAO » et qui est utilisée pour mesurer les distances dans l'univers, notamment le taux d'expansion de l'univers H et la distance angulaire D_A. A droite de la figure 6.4 est représentée la fonction de corrélation à deux points dans les directions parallèle et perpendiculaire à la ligne de visée (LOS pour Line of Sight). On distingue un anneau de rayon 100 h^{-1}Mpc qui correspond au pic BAO : une sur-densité de galaxies distantes de 100 h^{-1}Mpc. On observe également un écrasement de la fonction de corrélation le long de la LOS qui correspond à l'effet de distorsions dans l'espace des redshifts (RSD pour Redshift Space Distortions). Cet effet est relié à la mesure de la distance entre une galaxie et l'observateur à partir du redshift. La position de la galaxie notée s contient deux termes : une contribution due à l'expansion de l'univers (on parle de flot de Hubble) notée r et une autre contribution due à la composante le long de la LOS de la vitesse particulière de la galaxie notée v_z :

$$\mathbf{s} = \mathbf{r} + \frac{v_z(\mathbf{r})}{aH}\hat{\mathbf{z}} \tag{6.3}$$

où $\hat{\mathbf{z}}$ est le vecteur unitaire le long de la LOS. On appelle aussi s la position dans l'espace des redshifts et r la position dans l'espace réel. La prise en compte des vitesses propres des

galaxies introduit des déformations de la fonction de corrélation qui sont directement liées à la dynamique des structures qui s'attirent sous l'effet de l'attraction gravitationnelle. L'amplitude de l'effet de RSD est directement proportionnelle aux taux de croissance des structures cosmiques f. Le modèle actuel de la cosmologie, basé sur la relativité générale d'Einstein comme théorie de la gravitation, établit une prédiction théorique pour ce paramètre cosmologique :

$$f = \frac{d \ln D^{(+)}}{d \ln a} = \Omega_m^{\gamma} \tag{6.4}$$

où $\gamma = 0.55$ pour un univers ΛCDM plat. L'exploitation de l'effet de RSD permet donc de tester directement la théorie de la gravitation aux échelles cosmologiques.

6.4 L'échantillon de quasars de la DR14 du SDSS-IV eBOSS

La collaboration SDSS avaient déjà auparavant utilisé la technique des BAO et RSD avec des galaxies plus proches pour $z < 1$ pour mesurer les distances dans l'univers et la croissance des structures. Pour étendre les mesures faites jusqu'à présent grâce aux galaxies et être capable de sonder l'univers tel qu'il était il y a plus de 6 milliards d'années, il faut pouvoir observer des astres très brillants et très lointains. Les quasars sont les candidats idéaux pour réaliser la carte la plus aboutie à ce jour des grandes structures de l'univers. Il s'agit de galaxies actives avec un trou noir super-massif. L'incroyable luminosité des quasars est due à l'interaction entre le disque d'accrétion et le trou noir super-massif. Leurs processus de formation et les interactions avec la matière environnante font partie des phénomènes les plus violents de l'univers et sont encore méconnus. Le fait qu'ils se soient formés tôt dans l'histoire de l'univers permet de sonder une époque pratiquement inexplorée, lorsqu'il n'était pas encore dominé par une phase d'expansion accélérée. Des quasars lointains ($z > 2.5$) avaient déjà été observés par le programme précédent BOSS pour analyser la forêt Ly-α qui correspond à l'absorption de l'hydrogène par le milieu intergalactique. Avec eBOSS, l'objectif au contraire est d'utiliser les quasars entre $0.8 \leq z \leq 2.2$ comme traceurs direct du champ de matière.

Pour réaliser la carte la plus aboutie à ce jour des grandes structures de l'univers, le télescope SDSS, relié au multi-spectrographe eBOSS, a observé pendant deux ans un nombre sans précédent de quasars. Il a permis de mesurer précisément la position de plus de 149 000 quasars. La première phase de la stratégie d'observation correspond à un relevé photométrique de la zone étudiée obtenu avec la caméra de SDSS. Une analyse de ce relevé permet de définir quels sont les candidats quasars, appelés cibles, que le spectrographe va analyser afin d'obtenir leurs coordonnées le long de la ligne de visée, c'est-à-dire leur redshift. Plus de 300 plaques en aluminium ont alors été percées afin de correspondre parfaitement aux coordonnées des candidats pré-sélectionnés. Ces plaques sont placées au plan focal du télescope pour ne récupérer que la lumière émise par ces sources grâce à des fibres optiques.

6.4.1 Propriétés et traitement des effets de sélection

L'échantillon DR14 (Data Release 14) correspond à deux ans de prise de données dont les principales caractéristiques sont présentées dans le tableau 6.1. Il s'agit d'un échantillon peu dense avec une densité moyenne de $\bar{n} \sim 10^{-4}[h^{-1}\mathrm{Mpc}]^{-3}$ soit un ordre de grandeur plus faible que la densité moyenne des galaxies de BOSS.

Dans n'importe quel relevé du ciel réaliste, plusieurs raisons peuvent conduire à des variations de la densité de galaxies comme la géométrie du relevé, des effets observationnels,

TABLE 6.1 – Nombre de quasars entre $0.8 \leq z \leq 2.2$ pour l'échantillon eBOSS CORE et surface effective des pôles galactiques nord (NGC) et sud (SGC).

	NGC	SGC	Total
N_{quasar} $(0.8 \leq z \leq 2.2)$	89233	59426	148659
Surface effective (deg^2)	1214.6	898.3	2112.9

instrumentaux... Pour corriger ces effets, nous définissons des poids que nous appliquons aux données. Le schéma de pondération que nous adoptons permet de corriger les effets suivants :

— in-homogénéités dans la sélection de cibles dues à la photométrie : un poids est défini à partir de plusieurs observables comme la densité d'étoiles, la poussière galactique, la profondeur de la photométrie, ... Ce poids est appelé poids photométrique et est noté $w_{\rm photo}$.

— collisions de fibre : chaque plaque d'aluminium est percée de 1000 trous reliés au spectrographe par des fibres optiques. Le diamètre d'une fibre empêche deux quasars séparés de moins de 62" d'être observés. Certaines régions sont observés par plus d'une plaque et permettent de résoudre ces collisions. Lorsqu'un quasar est manqué, on rajoute un poids +1 au quasar le plus proche, ce poids est noté $w_{\rm cp}$.

— erreurs d'assignation de redshift : toutes les observations spectroscopiques ne permettent pas d'obtenir un redshift valide. La méthode utilisée jusqu'à présent constituait à rajouter un poids +1 au quasar le plus proche mais dans ma thèse, nous avons proposé une amélioration : on peut calculer une efficacité d'assignation de redshift en fonction de la position de la fibre dans le plan focal et donc donner définir un poids en fonction de l'inverse de cette efficacité. On note ce poids $w_{\rm focal}$

— poids pour minimiser la variance noté $w_{\rm FKP}$

Le poids total $\mathcal{W}_{\rm tot}$ est alors défini par :

$$\mathcal{W}_{\rm tot} = w_{\rm FKP} \cdot w_{\rm photo} \cdot w_{\rm cp} \cdot w_{\rm focal} \tag{6.5}$$

Le nombre effectif de quasars correspond alors à $N_{\rm quasars,eff} = \sum_i \mathcal{W}_{\rm tot,i}$ et on définit le redshift effectif $z_{\rm eff}$ par $z_{\rm eff} = \sum_i z_i \mathcal{W}_{\rm tot,i} \,/\, \sum_i W_{\rm tot,i} = 1.52$.

6.4.2 Incertitudes sur la détermination du redshift des quasars

La détermination du redshift procède de l'analyse de la position des raies spectrales. Un quasar possède un continuum qui correspond à l'émission thermique du disque d'accrétion qui entoure le trou noir super-massif ainsi que des raies d'émission étroites et larges associées à des régions situées autour du trou noir. La position de ces raies d'émission peut être décalée à cause d'éjection de matière lors de processus d'accrétion autour du trou noir, ces effets astrophysiques peuvent donc conduire à un biais systématique dans la détermination du redshift.

Au cours de ma thèse, j'ai étudié l'impact des incertitudes sur la détermination du redshift sur la fonction de corrélation des quasars. Plusieurs estimateurs du redshift sont disponibles dans le catalogue de quasars DR14, en particulier :

— le redshift de référence noté 'z' qui correspond au redshift déterminé par l'algorithme de traitement des spectres qui utilise plusieurs raies d'émission noté '$z_{\rm PL}$ ou au redshift déterminé par une inspection visuelle des spectres identifiés comme suspects par l'algorithme noté '$z_{\rm VI}$'.

— le redshift déterminé à partir de la position de la raie Magnesium II noté 'z_{MgII}' car pour des raisons astrophysiques et étant donné le rapport signal sur bruit des spectres eBOSS et l'intervalle en redshift sondé cet estimateur est considéré comme celui étant le moins biaisé.

Toutefois, tous les spectres ne présentent pas de raie MgII suffisamment prononcée pour fournir un z_{MgII} à tous les quasars de l'échantillon. Par conséquent, le redshift utilisé pour l'analyse reste z et pour avoir une estimation de la résolution en redshift, nous regardons le RMS et la forme de la distribution '$z - z_{\mathrm{MgII}}$'. La figure 6.5 montre la distribution des différences entre plusieurs estimateurs de redshift pour deux bins en redshift, la distribution '$z - z_{\mathrm{MgII}}$' est représentée en vert. On voit que par rapport à une distribution Gaussienne en noir pointillé, la distribution physique présente des queues non-Gaussiennes.

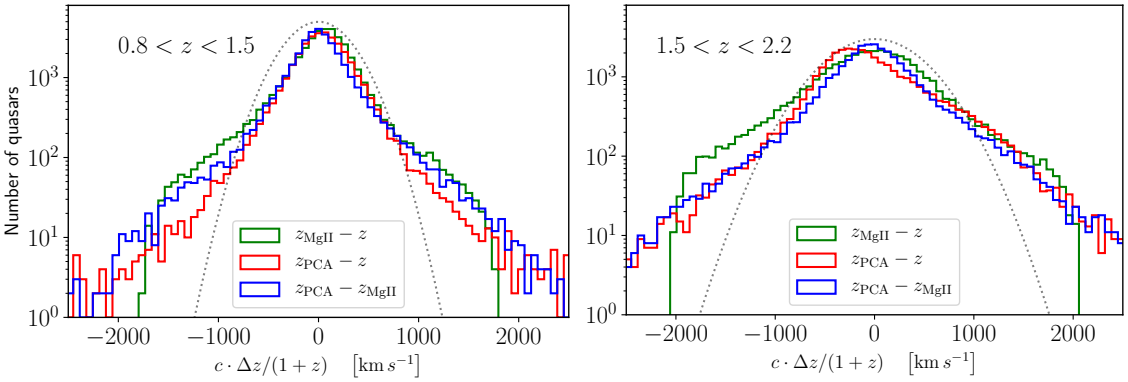


FIGURE 6.5 – Distribution des différences entre deux estimateurs du redshift pour deux domaines en redshift, $0.8 < z < 1.5$ à gauche et $1.5 < z < 2.2$ à droite.

6.5 Méthodologie de l'analyse 'clustering'

Nous commençons par construire des catalogues des grandes structures qui prennent en compte la géométrie angulaire et radiale du relevé. Nous utilisons l'estimateur de Landy-Szalay pour calculer la fonction de corrélation, il est défini par :

$$\hat{\xi}(s,\mu) = \frac{DD(s,\mu) - 2DR(s,\mu) + RR(s,\mu)}{RR(s,\mu)} \tag{6.6}$$

où s est la distance qui sépare deux quasars dans l'espace des redshift, μ est le cosinus de l'angle entre la paire et la LOS, DD correspond au nombre de paires dans le catalogue de données, RR au nombre de paires pour une distribution homogène de même densité moyenne obtenue en générant un catalogue de points répartis de manière uniforme dans le volume couvert par le relevé, et DR correspond au nombre de paires entre un quasar du catalogue de données et un point du catalogue aléatoire. Puis la fonction de corrélation est projetée sur la base des polynômes de Legendre :

$$\xi_l(s) = \frac{2l+1}{2} \sum_j \xi(s,\mu_j) P_l(\mu_j) d\mu \,, \tag{6.7}$$

où nous considérons uniquement $l = 0, 2, 4$ pour le monopole, le quadrupole et l'hexadecapole respectivement.

6.5.1 Analyse BAO

L'analyse dite BAO utilise uniquement les informations contenues dans le monopole.

Pour mesurer $\xi(r)$ et la position du pic BAO, nous avons besoin de transformer les coordonnées observationnelles des quasars en coordonnées cartésiennes (en h^{-1}Mpc). Pour effectuer cette transformation, il est nécessaire de définir une cosmologie fiducielle, c'est-à-dire de fixer les paramètres cosmologiques du modèle ΛCDM. Fixer ces paramètres cosmologiques fixe également la position théorique du pic BAO dans la fonction de corrélation et celle des oscillations dans le spectre de puissance. La position du pic est alors ajustée, tandis que la forme de la fonction de corrélation sans pic BAO (nommée *broad-band* en anglais) est reproduite à l'aide d'un modèle qui comporte généralement plusieurs paramètres de nuisance. Nous mesurons alors α, le rapport entre la position mesurée et la position théorique du pic BAO. Quand la statistique du relevé est trop faible pour obtenir une mesure des distances parallèles et perpendiculaires à la ligne de visée, il est préférable de réaliser une mesure isotrope pour former la distance $D_V(z)$:

$$D_{\rm V}(z) = \left[(1+z)^2 cz \frac{D_{\rm A}^2(z)}{H(z)}\right]^{\frac{1}{3}} \tag{6.8}$$

où $H(z)$ est le taux d'expansion de l'univers et D_A est la distance comobile angulaire. Une mesure isotrope permet alors de contraindre le rapport α_{iso} :

$$\alpha_{iso} = \frac{D_V(z)/r_d}{D_V^{fid}(z)/r_d^{fid}} \,. \tag{6.9}$$

La mesure de la position du pic BAO est très robuste, car il est peu probable qu'un effet systématique puisse générer un pic de corrélation. La plupart des effets systématiques seront absorbés par le *broadband*, et vont seulement élargir le pic et réduire la précision de la mesure. Cependant, la faiblesse de l'amplitude du pic nécessite un échantillon de données conséquent, donc les mesures sont limitées par la précision statistique.

Une analyse jointe du monopole de la fonction de corrélation et du spectre de puissance, sa transformée de Fourier ont permis de mesurer la position du pic BAO pour accéder à la distance D_V avec une précision de 3.8%. La mesure obtenue est compatible avec les mesures faites jusqu'à présent en utilisant les galaxies à $z < 1$ et les quasars avec leurs forêts Ly-α à $z > 2.5$.

6.5.2 Analyse "Full Shape"

L'analyse dite 'Full-Shape' utilise les informations contenues dans les trois multipoles. Contrairement à l'analyse BAO qui utilise un *broad band*, l'analyse Full-Shape nécessite de modéliser toute la fonction de corrélation, et ce pas uniquement aux échelles linéaires ($s > 80\,h^{-1}$Mpc) mais aussi aux échelles intermédiaires ($20-80\,h^{-1}$Mpc) où il est indispensable de prendre en compte des effets non-linéaires.

6.5.2.1 Modélisation de la fonction de corrélation à deux points

La construction d'une prédiction théorique de la fonction de corrélation à deux points requiert trois éléments :

— un modèle pour l'effet de RSD, i.e. le passage de l'espace réel à l'espace des redshifts

— une prédiction théorique pour le champ de densité et de vitesses, ainsi que leurs couplages aux échelles intermédiaires

— un modèle de biais pour relier les propriétés statistiques du champ de matière à celles des traceurs, ici les quasars de eBOSS qui évoluent dans des halos de matière noire dont la masse typique est de l'ordre de $10^{12}\,\mathrm{M}_{\odot}h^{-1}$

Le premier élément est traité de la manière suivante : la fonction de corrélation dans l'espace des redshifts $\xi(s,\mu)$ peut être modélisée comme étant la convolution de la fonction de corrélation dans l'espace réel $\xi(r)$ avec la densité de probabilité $\mathcal{G}$ des vitesses d'effondrement v_{12} de dispersion σ_{12} :

$$1+\xi^s(s,\mu) = \int d^3r_{\parallel}[1+\xi^r(r)]\mathcal{G}(s_{\parallel}-r_{\parallel},v_{12},\sigma_{12}) \tag{6.10}$$

Dans le modèle adopté, la densité de probabilité $\mathcal{G}$ peut être approximée par une Gaussienne centrée sur $\mu v_{12}(r)$ avec une dispersion qui dépend de l'échelle $\sigma_{12}(r,\mu)$ de sorte que $\mathcal{G}$ devient :

$$\mathcal{G} = \frac{1}{\sqrt{2\pi\sigma_{12}^2(r,\mu)}}\exp[\frac{(s_{\parallel}-r_{\parallel}-\mu v_{12})^2}{2\sigma_{12}^2(r,\mu)}] \tag{6.11}$$

Le modèle de RSD dépend donc de trois ingrédients, $\xi(r)$ relié à la variance du champ de densité, $v_{12}(r)$ relié au couplage densité-vitesse et $\sigma_{12}(r,\mu)$ relié à la variance du champ de vitesse. Les prédictions théoriques sont obtenues en utilisant un des modèles les plus récents qui prend en compte l'évolution non-linéaire des champs de matière et de vitesse en se basant sur la théorie des perturbations dans le formalisme Lagrangien (CLPT pour Convolution Lagrangian Perturbation Theory). Chaque ingrédient est prédit de la manière suivante :

$$A = A_0 + \langle F'\rangle A_{10} + \langle F''\rangle A_{01} + \langle F'\rangle^2 A_{20} + \langle F''\rangle^2 A_{02} + \langle F'\rangle\langle F''\rangle A_{11} \tag{6.12}$$

où chaque coefficient A_{ij} coefficient est calculé avec le code public CLPT [2]. Les termes A_{10} et A_{01} correspondent) des corrections du premier ordre ; A_{20}, A_{02} et A_{11} des corrections du deuxième ordre.

Les paramètres F' et F'' correspondent aux paramètres du modèle de biais qui permettent d'obtenir les propriétés statistiques des quasars à partir de celles des halos de matière noire. Le modèle de biais adopté est un modèle non-linéaire et local dans le formalisme Lagrangien.

6.5.2.2 Étude des effets systématiques

Les effets systématiques pouvant affecter la mesure des paramètres cosmologiques peuvent se classer en 2 catégories :

— Effets systématiques liés à la modélisation de la fonction de corrélation à deux points : le modèle avait déjà été utilisé pour une analyse antérieure sur l'échantillon de galaxies proches obtenu avec BOSS mais il n'a jamais été testé pour un échantillon de quasars beaucoup plus lointains. Par conséquent, j'ai étudié les performances et les conditions d'utilisation du modèle à partir de simulations numériques d'univers. J'ai participé à la génération d'une centaine de catalogues fictifs qui reproduisent notre échantillon de données et pour lesquels nous commençons les paramètres cosmologiques injectés. Les effets systématiques liés à notre modèle correspondent à 40% de la précision statistique où plus de la moitié provient des incertitudes sur le redshift. Pour étudier leurs impacts, j'ai généré des catalogues fictifs avec différentes distributions d'incertitudes sur le redshift : pas d'incertitudes, une distribution Gaussienne

2. CLPT

d'après la courbe en pointillé noir de la figure 6.5 et une distribution physique d'après la courbe verte de la même figure. Les multipoles pour ces trois configurations sont montrés sur la figure de gauche 6.6 où l'on voit que le quadrupole est le plus impacté par la résolution en redshift, surtout pour des échelles $s < 40\,h^{-1}$Mpc. Il est possible de mesurer les incertitudes sur le redshift en rajoutant un paramètre de dispersion de vitesse supposé constant dans le modèle noté $\sigma_{\rm tot}$ qui inclut aussi les effets de dispersion des vitesses à petites échelles au sein d'un halo.

— Effets systématiques liés aux conditions d'observation et à l'instrument : j'ai testé différents schémas de pondérations sur 1 000 catalogues fictifs reproduisant nos données mais avec une résolution moins bonne que les catalogues précédents basés sur des simulations à N-corps. Ces catalogues sont aussi utilisés pour déterminer la matrice de covariance à partir de laquelle nous estimons les erreurs statistiques. J'ai montré que le nouveau schéma de pondération $\mathcal{W}_{\rm tot} = \mathcal{W}_{{\rm focal}-\mu}$ que nous avons proposé pour l'analyse Full-Shape des quasars permettait de retrouver les mulltipoles avec une meilleure précision comparé à l'approche classique $\mathcal{W}_{\rm noz}$, notamment le quadrupole comme le montre la figure de droite 6.6. Grâce à l'amélioration de la définition des poids, nous n'avons pas de systématiques observationnelles résiduelles.

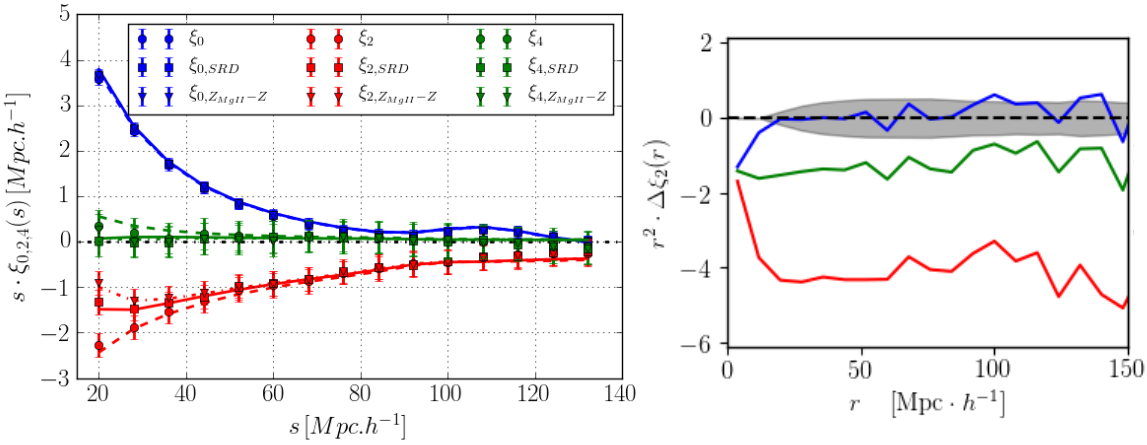


FIGURE 6.6 – Gauche : Monopole (bleu), quadrupole (rouge) et hexadecapole (vert) de la fonction de corrélation pour 3 configurations de catalogues fictifs générés à partir d'une simulation numérique à N-corps : sans incertitude sur le redshift (pointillé), distribution Gaussienne (plein), distribution physique '$z - z_{\rm MgII}$' (tiret). Droite : Effet des différents schémas de pondération sur le quadrupole : en rouge lorsqu'on applique un poids +1 au quasar le plus proche pour corriger des erreurs d'assignation de redshift, en vert $\mathcal{W}_{\rm focal}$ lorsqu'on applique un poids en fonction de l'inverse de l'efficacité d'assignation de redshift, en bleu $\mathcal{W}_{{\rm focal}-\mu}$ on rajoute un traitement spécial pour les collisions de fibre où l'on retire les paires dans la région proche de la LOS.

La figure 6.7 montre les multipoles de la fonction de corrélation des données quasars DR14 où les barres d'erreur statistique sont obtenues à partir de la matrice de covariance des 1 000 générations fictives. Les courbes en tiret correspondent aux prédictions théoriques obtenues à partir du modèle décrit précédemment qui dépend de 5 paramètres : 1 paramètre de biais linéaire F' ; 3 paramètres cosmologiques : décalage parallèle et perpendiculaire à la ligne de visée du pic BAO $\alpha_\parallel$, $\alpha_\perp$ et le taux de croissance des structures f ; 1 paramètre qui contient les incertitudes sur le redshift $\sigma_{\rm tot}$. Les paramètres $\alpha_\parallel$ et $\alpha_\perp$ sont

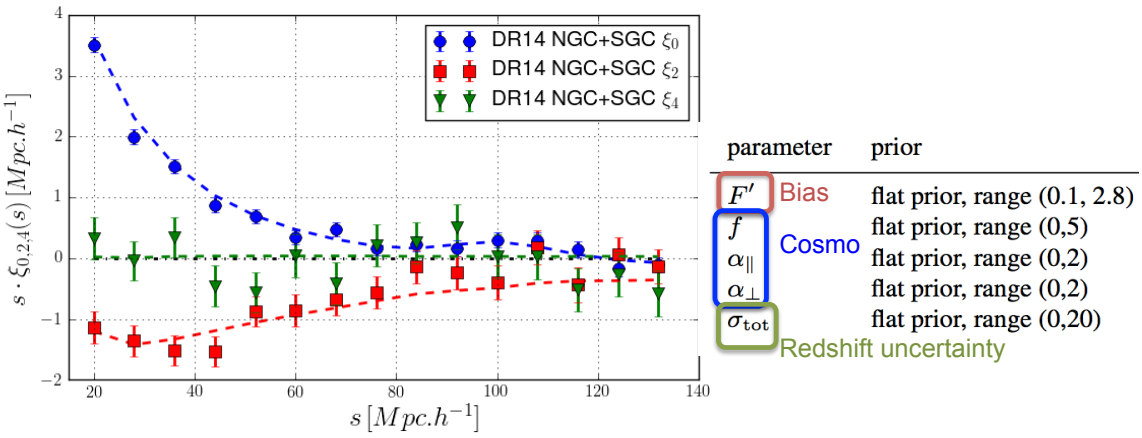


FIGURE 6.7 – Monopole (bleu), quadrupole (rouge) et hexadecapole (vert) de la fonction de corrélation des quasars de DR14. Les courbes en tiret correspondent aux prédictions théoriques du modèle utilisé qui dépendent de 5 paramètres définis dans le tableau à droite avec leurs intervalles de valeurs respectifs.

TABLE 6.2 – Resultats des ajustements des données avec les incertitudes statistique et systématique.

paramètre	best-fit	statistique	systématique
$b\sigma_8$	1.038	$^{+0.060}_{-0.057}$	–
$f\sigma_8$	0.426	$^{+0.070}_{-0.070}$	0.033
$\alpha_{\rm par}$	1.012	$^{+0.071}_{-0.064}$	0.038
$\alpha_{\rm perp}$	1.031	$^{+0.050}_{-0.048}$	0.006
$\sigma_{\rm tot}$	5.94	$^{+1.19}_{-1.40}$	–

reliés au taux d'expansion de l'univers $H(z)$ et à la distance comobile angulaire D_A par :

$$\alpha_\parallel = \frac{H^{\rm fid}(z) r_d^{\rm fid}}{H(z) r_s} \tag{6.13}$$

$$\alpha_\perp = \frac{D_A(z) r_d^{\rm fid}}{D_A^{\rm fid}(z) r_s} \tag{6.14}$$

Les résultats de l'ajustement des données DR14 quasars avec le modèle décrit précédemment sont présentés dans le tableau 6.2, ainsi que les incertitudes statistique et systématique où seule une incertitude liée à la modélisation est prise en compte.

6.6 Résultats et implications cosmologiques

Les résultats de l'analyse BAO ont été publiés dans Ata et al. (2017) [3], il s'agit de la première détection du BAO entre $1 < z < 2$ réalisée avec un échantillon de quasars

3. M. Ata et al., *The clustering of the SDSS- IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : First measurement of Baryon Acoustic Oscillations between redshift 0.8 and 2.2*, arXiv:1705.06373.

TABLE 6.3 – Constraints on the BAO distances from BAO-only analyses using various large-scale structure surveys.

survey	reference	z	D_V	D_M	H
6dFGS BAO	Beutler et al. 2011	0.106	4.5%	–	–
WiggleZ BAO	Kazin et al. 2014	0.44	4.8%	–	–
		0.60	4.5%	–	–
		0.73	3.4%	–	–
WiggleZ FS	Blake et al. 2012	0.44	–	9.4%	9.4%
		0.60	–	6.9%	6.9%
		0.73	–	7.0%	7.2%
SDSS MGS BAO	Ross et al. 2015	0.15	3.8%	–	–
SDSS BOSS galaxies BAO+FS	Alam et al. 2016	0.38	1.09%	1.5%	2.4%
	Alam et al. 2016	0.51	1.01%	1.4%	2.2%
	Alam et al. 2016	0.61	1.03%	2.8%	2.2%
SDSS BOSS Ly-α BAO	Bautista et al. 2017	2.40	–	5.8%	3.4%
	du Mas des Bourboux et al. 2017	2.33	–	3.9%	2.8%
	combinaison	2.40	–	3.0%	2.0%
SDSS eBOSS LRG BAO	Bautista et al. 2018	0.72	2.6%	–	–
SDSS eBOSS quasars BAO	Ata et al. 2017	1.52	3.8%	–	–
SDSS eBOSS quasars FS	Zarrouk et al. 2018	1.52	4.1%	5.5%	7.9%

suffisamment important. Les résultats de l'analyse Full-Shape utilisant la fonction de corrélation ont été publiés dans Zarrouk et al. (2018) [4].

6.6.1 Distances cosmiques

Le tableau 6.3 résume les mesures des distances cosmiques à partir des analyses BAO et Full-Shape en utilisant des données de différents relevés des grandes structures. L'analyse de l'échantillon quasars DR14 apporte les contraintes suivantes sur le taux d'expansion de l'univers $H(z_{\rm eff}) = 159^{+12}_{-13}(r_s^{\rm fid}/r_s){\rm km.s^{-1}.Mpc^{-1}}$, et la distance angulaire $D_A(z_{\rm eff}) = 1850^{+90}_{-115}\,(r_s/r_s^{\rm fid}){\rm Mpc}$ où r_s correspond à l'horizon sonore au moment de la recombinaison et $r_s^{\rm fid}$ est sa valeur dans le modèle de cosmologie fiducielle adoptée pour l'analyse. Les résultas obtenus avec l'analyse BOSS DR12 fournissent les meilleures contraintes avec un précision de l'ordre de 1-3% aux redshifts 0.38, 0.51, 0.61 avec des galaxies et au redshift $z = 2.4$ avec les forêts Ly-α. Nos mesures de distance obtenues avec le monopole uniquement (analyse BAO) et avec toute la fonction de corrélation (analyse Full-Shape) sont présentées dans le tableau.

La figure 6.8 présente les mesures de distances cosmiques comparées à la prédiction donnée par le modèle Λ-CDM avec les paramètres de *Planck* 2015. Nos mesures avec l'analyse BAO et Full-Shape sont compatibles entre elles et sont compatibles avec les mesures des autres analyses. Elles sont en accord avec l'expansion de l'univers décrite par le modèle standard.

4. P. Zarrouk, E. Burtin, et al., *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample : measurement of the growth rate of structure from the anisotropic correlation function between redshift 0.8 and 2.2*, 477, 1639 (2018), doi:10.1093/mnras/sty506.

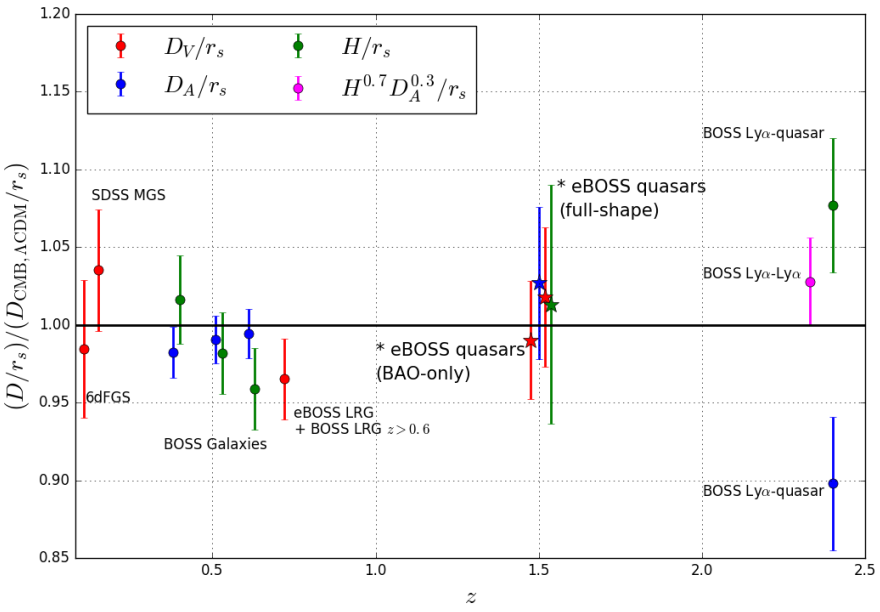


FIGURE 6.8 – Evolution des distances BAO avec le redshfit comparée à la prédiction du modèle Λ-CDM pour un univers plat avec les paramètres de *Planck*. Les résultats de cette thèse utilisant les données eBOSS quasars sont représentés par une étoile *.

6.6.2 Croissance des structures

L'analyse des anisotropies de la fonction de corrélation des quasars DR14 eBOSS a conduit à mesurer le taux de croissance des structures $f\sigma_8(z_{eff} = 1.52) = 0.426 \pm 0.079$ qui est représenté sur la figure 6.9. Ce résultat est obtenu en réalisant un ajustement des multipoles de Legendre $l = 0, 2, 4$, la barre d'erreur contient le incertitudes statistique et systématiques (liées au modèle de RSD) ajoutées en quadrature. Notre mesure est compatible avec les mesures précédentes et est en accord avec la prédiction du modèle ΛCDM basé sur la relativité générale et confirme donc la validité de la théorie pour un intervalle en redshift intermédiaire ($1 < z < 2$).

Notre mesure a été réalisée en laissant varier le modèle cosmologique que nous devons supposer pour transformer les coordonnées galactiques en coordonnées cartésiennes, ce qui revient à réaliser un ajustement des données avec les paramètres $f\sigma_8$, H et D_A. Ce n'est pas le cas de toutes les mesures effectuées, celles en tiret ont été obtenues pour un modèle cosmologique fixée ($H = H^{fid}$, $D_A = D_A^{fid}$).

Le tableau 6.4 résume les mesures de $f\sigma_8$ et précise celles qui ont été effectuées en marginalisant sur H et D_A donc en laissant varier le modèle cosmologie. Lorsque la cosmologie est fixée, nous obtenons une mesure de $f\sigma_8$ à 11%. Toutefois, ilest important de souligner que si un ajustement joint de $f\sigma_8$, H et D_A n'est pas réalisée, l'incertitude sur $f\sigma_8$ ne reflète pas la marginalisation sur le modèle cosmologique et la mesure ne peut donc pas être utilisée pour contraindre des scénarios de gravité modifiée.

6.6.3 Contraintes sur ΛCDM et au-delà

En utilisant notre résultat de l'analyse BAO combiné avec d'autres mesures de BAO indépendantes, les contraintes actuelles favorisent un modèle d'univers avec une densité d'énergie noire sous la forme d'une constante cosmologique non nulle avec $\Omega_\Lambda > 0$ à 6.6σ

TABLE 6.4 – Contraintes sur $f\sigma_8$, D_A and H obtenues à partir d'analyses Full-Shape provenant de différents relevés des grandes structures.

survey	reference	z	$f\sigma_8$	cosmologie fixée
2dFGRS	Percival et al. 2004	0.17	13%	oui
6dFGS	Beutler et aL. 2012	0.067	13%	oui
WiggleZ	Blake et al. 2012	0.44	19.4%	non
		0.60	16.2%	non
		0.73	16.5%	non
VVDS	Guzzo et al. 2008	0.77	36.7%	oui
VIPERS	Pezzotta et al. 2017	0.60	21.8%	oui
		0.86	27.5%	oui
Subaru FastSound	Okumura et al. 2016	1.4	25%	oui
SDSS MGS	Howlett et al. 2015	0.15	30.6%	non
SDSS BOSS galaxies FS	Alam et al. 2012	0.38	9.5%	non
		0.51	8.7%	non
		0.61	8.8%	non
SDSS BOSS galaxies BAO+FS	Alam et al. 2012	0.38	9.2%	non
	Alam et al. 2012	0.51	8.3%	non
	Alam et al. 2012	0.61	8.0%	non
SDSS eBOSS quasars	Zarrouk et al. 2018	1.52	18.1%	non

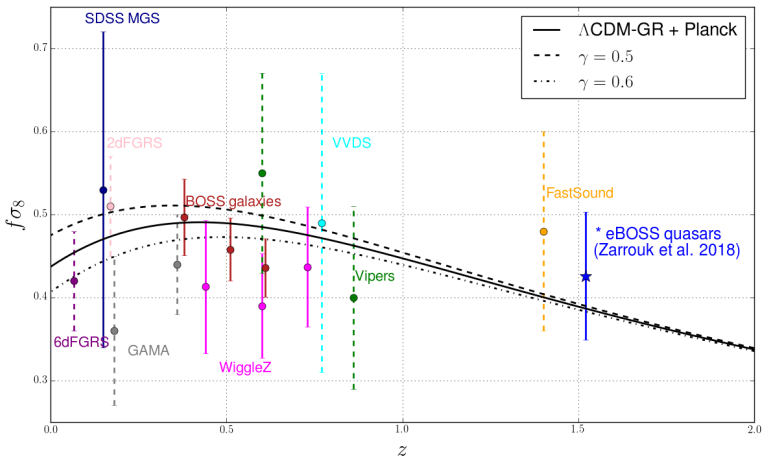


FIGURE 6.9 – Mesures de $f\sigma_8(z)$ en fonction du redshift comparées à la prédiction du modèle Λ-CDM pour un univers plat avec les paramètres de *Planck*. Le résultat de ce travail de thèse avec les quasars eBOSS DR14 est représenté par une étoile * et la barre d'erreur contient les incertitudes statistique et systématiques.

de significance.

En combinant les mesures effectuées dans cette thèse à partir de l'analyse Full-Shape sur les distances avec d'autres mesures indépendantes (BOSS galaxies et Ly-α + *Planck*), nous obtenons les contraintes suivantes sur le modèle ΛCDM :

$$\Omega_m = 0.3094^{+0.0076}_{-0.0080} \tag{6.15}$$

$$\Omega_\Lambda = 0.697^{+0.035}_{-0.032} \tag{6.16}$$

$$\Omega_k = -0.007 \pm 0.030 \tag{6.17}$$

Les résultats sont en accord avec un univers dominé par une constante cosmologique et sans courbure.

Concernant les extensions au modèle ΛCDM, nous considérons :

— une extension minimale à la relativité général consiste à faire varier l'indice de croissance γ où $f\sigma_8 = \Omega_m^{\gamma=0.55}$. En combinant les résultats de ce travail sur l'analyse Full-Shape des quasars eBOSS avec *Planck*, nous trouvons : $\gamma = -0.2 \pm 1.2$. Le manque de précision provient du fait qu'au redshift des quasars, Ω_m est proche de 1 car l'univers est dominé par la matière et pas encore par sa phase d'expansion accélérée.

— un modèle de gravité modifiée en particulier a été étudié dans mon groupe, il s'agit du Galileon qui rajoute un champ scalaire supplémentaire qui peut se coupler à la matière. Le modèle complet présente 5 Lagrangians supplémentaires par rapport à la RG. Un ajustement de ce modèle sur les données *Planck*, l'analyse BAO des galaxies BOSS et les supernovae JLA a été réalisé, et est exclu par la réionisation. De plus, le modèle a été exclu par la détection jointe d'une onde gravitationnelle et d'un sursaut γ qui indique une vitesse de propagation des ondes gravitationnelles et des ondes électromagnétiques très proche, ce qui n'est pas le cas dans le cadre du modèle complet du Galileon. Dans mon travail j'ai regardé a posteriori la prédiction de $f\sigma_8$ faite par ce modèle avec les paramètres résultant de l'ajustement. La prédiction est en accord avec ma mesure mais présente quelques tensions avec les mesures de croissance des structures à plus bas redshift.

— une autre catégorie de modèles de gravité modifiée correspond aux modèles scalaire-tenseur, dont le plus connu est le modèle $f(R)$ avec un degré de liberté supplémentaire. L'ajout de champ supplémentaire modifie les équations de Poisson et introduit des dépendances en fonction de l'échelle et du temps qui peuvent être exprimées avec la paramétrisation BZ. Notre mesure seule n'a pas encore la précision requise pour contraindre fortement l'espace des paramètres de ces modèles aux redshifts intermédiaires $1 < z < 2$ mais en combinant les mesures à différents redshifts, on contraint de mieux en mieux l'histoire de la croissance des structures, et donc la théorie de la gravitation sous-jacente.

Ce travail de thèse constitue une première étude menée avec les données de quasars de eBOSS et sera utilisé pour l'analyse de l'échantillon final à la fin 2019 où l'on attend une amélioration de la précision statistique d'un facteur 2. La première détection du pic BAO dans un échantillon de quasars à un redshift intermédiaire ($1 < z < 2$) combiné aux autres mesures à plus bas redshift avec les galaxies BOSS et à plus grand redshift avec les Ly-α a permis de confirmer le scénario d'un modèle d'univers avec constante cosmologique et sans courbure. Les mesures de croissance des structures à grand redshift, comme la première mesure que j'ai effectuée avec l'échantillon quasars de eBOSS vont ainsi apporter des nouvelles contraintes sur les modèles de gravité alternative, des contraintes qui sont

complémentaires de celles qui sont en train d'être apportées par les ondes gravitationnelles et qui s'annoncent aussi très prometteuses. Associé à BOSS, eBOSS ouvrira la voie pour les futurs programmes d'observation, comme le télescope au sol DESI et le satellite *Euclid*. Ces deux programmes sonderont intensivement l'époque de l'univers entre $1 < z < 2$ en observant plusieurs millions de spectres, ce qui permettra d'améliorer d'un ordre de grandeur au moins les contraintes actuelles sur les paramètres cosmologiques.

université PARIS-SACLAY | ÉCOLE DOCTORALE PHENIICS

Titre : Analyse des corrélations spatiales des quasars et implications cosmologiques avec le multi-spectrographe SDSS-IV eBOSS

Mots-clés : *énergie noire, oscillations acoustiques de baryons, distorsions dans l'espace des redshifts, taux de croissance des structures, gravité modifiée, relevé spectroscopique*

Résumé : Le modèle ΛCDM de la cosmologie repose sur l'existence d'une composante exotique, appelée énergie noire, pour expliquer l'accélération tardive de l'expansion de l'univers à $z < 0.7$. Des alternatives à cette constante cosmologique proposent de modifier la théorie de la gravitation basée sur la relativité générale aux échelles cosmologiques. Depuis l'automne 2014, le multi-spectrographe SDSS-eBOSS effectue un relevé de quasars dans un domaine en redshift peu exploré entre $0.8 \leq z \leq 2.2$ dont l'un des objectifs majeurs est d'étendre les contraintes sur la nature de l'énergie noire et de tester la validité de la théorie de la relativité générale à plus haut redshift en utilisant les quasars comme traceurs de la matière. Dans cette thèse, nous mesurons et analysons la fonction de corrélation à deux points de l'échantillon de quasars obtenu après deux ans d'observation d'eBOSS pour contraindre les distances cosmiques, à savoir la distance angulaire D_A et le taux d'expansion H, ainsi que le taux de croissance des structures $f\sigma_8$ à un redshift effectif $z_{\rm eff} = 1.52$. Nous commençons par construire des catalogues des grandes structures qui prennent en compte la géométrie angulaire et radiale du relevé. Puis pour obtenir des contraintes robustes, nous identifions plusieurs sources d'effets systématiques, en particulier ceux liés à la modélisation et aux observations sont étudiées avec des « mock catalogues » dédiés qui correspondent à des réalisations fictives de l'échantillon de quasars eBOSS. Les paramètres cosmologiques de ces catalogues fictifs étant connus, ils sont utilisés comme référence pour tester notre procédure d'analyse. Les résultats de ce travail sur l'évolution des distances cosmiques sont compatibles avec les prédictions du modèle ΛCDM utilisant les paramètres de *Planck* et basé sur l'existence d'une constante cosmologique. La mesure du taux de croissance des structures est compatible avec la prédiction de ce modèle basée sur la relativité générale, ce qui étend ainsi la validité de la théorie aux échelles cosmologiques à grand redshift. Nous utilisons également notre mesure pour mettre à jour les contraintes sur les modèles d'extensions à ΛCDM et sur les scénarios de gravité modifiée. Ce travail de thèse constitue une première étude menée avec les données de quasars d'eBOSS et sera utilisé pour l'analyse de l'échantillon final à la fin 2019 où l'on attend une amélioration de la précision statistique d'un facteur 2. Associé à BOSS, eBOSS ouvrira la voie pour les futurs programmes d'observation, comme le télescope au sol DESI et le satellite *Euclid*. Ces deux programmes sonderont intensivement l'époque de l'univers entre $1 < z < 2$ en observant plusieurs millions de spectres, ce qui permettra d'améliorer d'un ordre de grandeur au moins les contraintes actuelles sur les paramètres cosmologiques.

Title : Clustering analysis in configuration space and cosmological implications of the SDSS-IV eBOSS quasar sample

Keywords : *dark energy, baryon acoustic oscillations, redshift space distortions, growth rate of structure, modified gravity, spectroscopic survey*

Abstract : The ΛCDM model of cosmology assumes the existence of an exotic component, called dark energy, to explain the late-time acceleration of the expansion of the universe at redshift $z < 0.7$. Alternative scenarios to this cosmological constant suggest to modify the theory of gravitation based on general relativity at cosmological scales. Since fall 2014, the SDSS-IV eBOSS multi-object spectrograph has undertaken a survey of quasars in the almost unexplored redshift range $0.8 \leq z \leq 2.2$ with the key science goal to complement the constraints on dark energy and extend the test of general relativity at higher redshifts by using quasars as direct tracers of the matter field. In this thesis work, we measure and analyse the two-point correlation function of the two-year data taking of eBOSS quasar sample to constrain the cosmic distances, i.e. the angular diameter distance D_A and the expansion rate H, and the growth rate of structure $f\sigma_8$ at an effective redshift $z_{\rm eff} = 1.52$. First, we build large-scale structure catalogues that account for the angular and radial incompleteness of the survey. Then to obtain robust results, we investigate several potential systematics, in particular modeling and observational systematics are studied using dedicated mock catalogs which are fictional realizations of the data sample. These mocks are created with known cosmological parameters such that they are used as a benchmark to test the analysis pipeline. The results on the evolution of distances are consistent with the predictions for ΛCDM with *Planck* parameters assuming a cosmological constant. The measurement of the growth of structure is consistent with general relativity and hence extends its validity to higher redshift. We also provide updated constraints on extensions of ΛCDM and models of modified gravity. This study is a first use of eBOSS quasars as tracers of the matter field and will be included in the analysis of the final eBOSS sample at the end of 2019 with an expected improvement on the statistical precision of a factor 2. Together with BOSS, eBOSS will pave the way for future programs such as the ground-based Dark Energy Spectroscopic Instrument (DESI) and the space-based mission *Euclid*. Both programs will extensively probe the intermediate redshift range $1 < z < 2$ with millions of spectra, improving the cosmological constraints by an order of magnitude with respect to current measurements.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190, Saint-Aubin, France